

Spécifications de l'interface d'annotation de CADERIGE

Version 8.2 du 25/02/03

XANTED XML ANnotation Text Editor (nom à définir)

(Diffusion restreinte au consortium Caderige)

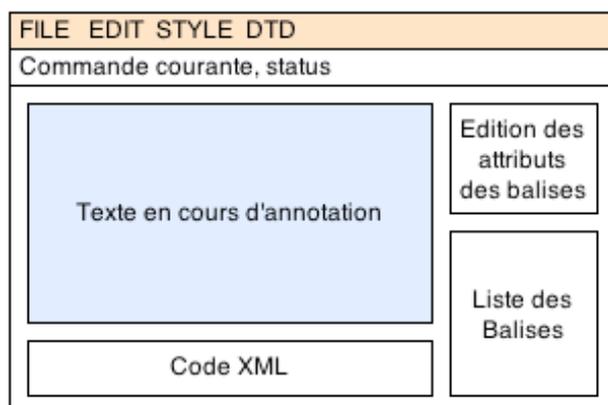
Objectifs

Le but est de concevoir un éditeur permettant d'annoter interactivement tout ou partie d'un texte à l'aide de balises XML décrites dans une DTD. 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 restant dans une optique « traitement de texte ». Les zones de textes balisés sont visualisées à l'aide feuilles styles librement éditables. Le logiciel est écrite en JAVA 1.3 et est portable sur Linux, MacOS X et Windows.

Présentation générale

L'interface est composée de 5 parties :

- MN : Le menu principal
- IC : Icône des commandes d'édition principales et éléments d'information
- TE : Editeur contenant le texte du document en cours d'annotation
- XM : Le code XML correspondant au texte en cours d'annotation
- EA : Editeur permettant de modifier les attributs de la balise courante
- LB : Liste des balises que l'on peut appliquer au texte



Afin de simplifier l'utilisation, le fonctionnement général de l'éditeur doit rester aussi proche que possible de celui d'un traitement de texte. Une fois le document de travail chargé, l'utilisateur sélectionne à la souris les zones sur lesquelles il veut appliquer une balise XML. Cette zone balisée est ensuite matérialisée dans l'éditeur à l'aide d'un ensemble d'attributs graphiques (couleurs, forme, police, taille ...) qui sont décrits dans des « feuilles de styles ». Initialement le texte que l'on charge dans l'éditeur est dépourvu de balise, c'est l'annotateur qui les introduit une à une dans l'ordre qu'il souhaite.

Point d'insertion

La position du curseur détermine la « balise courante » c'est à dire la balise qui englobe le bloc de texte dans lequel est placé le curseur. La partie EA indique à tout moment le nom de cette balise, la liste des attributs qui lui sont associés et leur valeur actuelle, soit sous la forme de champ de texte, soit sous la forme de menu lorsque le type est « énuméré ».

Dans la mesure où certaines balises peuvent avoir des positions de débuts et fin identiques dans le document annoté (cas de balises « imbriquée »), il est nécessaire que l'utilisateur puisse *facilement et rapidement* modifier le point d'insertion du texte. Cela s'effectue en sélectionnant (simple clic) dans la zone XML de l'éditeur l'étiquette dans laquelle il veut se placer. Une succession de « simples clics » déplace cycliquement le curseur d'insertion en début et fin de la zone. Un double-clic sélectionne l'ensemble du texte contenu dans la balise ce qui permet de visualiser et manipuler les blocs annotés.

Contrôle du balisage

Bien que l'on puisse valider « a posteriori » le texte annoté avec la DTD certaines situations peuvent être évitées « a priori » notamment le balisage avec recouvrement qui conduit à un code XML non valide. Par exemple :

- 1) Tag A : « un exemple de phrase » : un exemple <A> de phrase
- 2) Tag B : « un exemple de phrase » : un exemple <A> de phrase

La solution retenue consiste à n'activer à tout instant dans la zone LB de l'éditeur que les balises qui sont « potentiellement » acceptables, c-à-d celles qui sont « plus spécifiques » dans la DTD que la balise « courante » qui est implicitement indiquée par la position du curseur ou celle de la zone sélectionnée. Pour chaque balise, cette liste de dépendances peut être facilement calculée une fois pour toute lors de la lecture de la DTD. Cette notion de balise « plus spécifique » doit être implémentée de 2 manières qui sont sélectionnables via une option dans la sélection des « préférences » :

- 1) « Stricte », on n'accepte d'insérer que les balises immédiatement plus spécifiques que la balise courante (c'est l'implémentation actuelle). Cette approche permet à l'utilisateur novice de mieux comprendre la structure de la DTD
- 2) « Large » on accepte à tout instant d'insérer toutes les balises qui sont plus spécifiques que la balise courante. Cette seconde interprétation est cohérente avec le fonctionnement attendu de l'éditeur, qui vise à laisser l'utilisateur (non novice) libre de baliser le texte comme il l'entend ...

Toutefois le contrôle du balisage à la saisie ne détecte pas l'ensemble des problèmes, car la détermination des balises valides ne s'effectue que par rapport au début et la fin de la zone sélectionnée. Dès lors, l'utilisateur peut imbriquer des balises même si, dans la DTD, cette imbrication n'est pas autorisée. Par exemple :

- 1) « un exemple d'imbrication » : un exe<A> mple d'imrication
- 2) « un exemple d'imbrication » : un <A> exe<A> mple d'imrication

La détection de ses problèmes s'effectue actuellement via le vérificateur de DTD (la version « non-strict »). Il serait idéal que l'ensemble des problèmes soit détecté lors de la saisie, mais un appel systématique au validateur risque d'être trop coûteux ...

Modification du balisage

Dans la mesure où l'annotation est une tâche « d'analyse », l'utilisateur est conduit à modifier fréquemment les frontières des zones balisées. Cependant, la nature hiérarchique des DTD et la présence d'attributs valués, rendent difficiles la mise place de mécanismes de correction généraux et non ambigus qui fonctionnent par simple sélection de zones et (ré)application de balises comme on les trouve dans les traitements de textes¹. Par ailleurs, n'offrir comme possibilité d'édition que la simple suppression/réinsertion d'étiquettes, est trop limité car dans le cas où les balises comportent des attributs il faut ressaisir tous les informations à chaque modification.

On propose quatre opérations de modification de « limites » pour la balise courante qui permettent une édition relativement simple :

- EG : Extension de la balise courante vers la gauche
- ED : Extension de la balise courante vers la droite
- RG : Réduction de la balise courante vers la gauche
- RD : Réduction de la balise courante vers la droite

Ces opérations seraient « a priori² » exécutables tant qu'elles ne rentrent pas en « conflit » avec les autres balises c'est à dire tant qu'elle ne tendent pas à les « déplacer ». En cas de problème la commande n'est pas appliquée avec (si possible) un « feedback » sonore (bip). Les opérations sont applicables dans trois contextes différents :

- Sur la balise définie par le point d'insertion :
 - Le décalage est effectué « mot par mot »
- Sur la balise définie par le point d'insertion avec la touche CTRL enfoncée :
 - Le décalage est effectué « caractère par caractère »
- Sur une sélection
 - Dans ce cas, la balise courante et le décalage à effectuer sont déterminés par les positions de début et de fin de la zone sélectionnée.

Opération	Début de zone	Fin de zone
ED	Dans la balise à étendre	Nouvelle position de fin
EG	Nouvelle position de début	Dans la balise à étendre
RD	Position \leq au début de la balise	Nouvelle position de fin
RG	Nouvelle position de début	Position \geq à la fin de la balise

Par exemple dans le cas d'une sélection avec la commande ED :

Avant ED: « un **exemple** d'extension » un <A>exemple d'extension
Sélection: « un **exemple** d'extension »
Après ED: « un **exemple** d'imbrication » un <A>exemple d'extension

Pour les commandes RD (et RG) le point de début (et fin) de zone n'a pas d'importance puisqu'une seule information suffit pour déterminer la balise et la modification.

¹ Par exemple, dans un traitement de texte pour faire passer une zone de texte plus étendue en caractères gras, l'utilisateur sélectionne la nouvelle zone puis réapplique le passage en gras. En XML, une telle commande est ambiguë si les étiquettes courantes peuvent être imbriquées, car on ne sait pas si l'utilisateur veut étendre la taille de la zone ou créer un nouveau niveau dans la hiérarchie.

² Il faudrait regarder s'il n'est pas possible dans certains cas de « propager » automatiquement des décalages, notamment lorsque les balises imbriquées ne sont pas liées hiérarchiquement.

Barre de menu

Le contenu de chacun des items de la barre de menu de l'éditeur est le suivant. Chacune des fonctions associées est décrite dans la suite du document. Quelques points généraux sur le nommage de l'intitulé d'une commande dans le menu :

- Une commande commence toujours par une majuscule
- Lorsque la commande fait apparaître une BD, elle se termine pas « ... »
- Les commandes sont regroupées en séries homogènes séparées par un séparateur

File

About ...	: Affichage du « splash screen » pour identifier la version du logiciel
New document	: Création d'un nouveau document textuel vide
Load text document ...	: Chargement d'un nouveau document textuel
Load annotated document ...	: Chargement d'un document en cours d'annotation
Save annotated document ...	: Sauvegarde du document courant et de l'environnement de travail
Export HTML document ...	: Création d'un fichier HTML correspondant au texte annoté
Export RTF document ...	: Création d'un fichier RTF correspondant au texte annoté
Print ...	: Impression du document courant
Quit	: Sortie du logiciel avec sauvegarde

Edit

Cut text	: Coupe le texte « brut » sélectionné (sans les balises)
Copy text	: Copie le texte « brut » sélectionné (sans les balises)
Paste text	: Colle le texte « brut » dans le presse-papier
Cut XML block	: Coupe le bloc XML courant
Copy XML block	: Copie le bloc XML courant
Paste XML block	: Copie le bloc XML dans le presse papier si c'est possible
Delete current tag	: Supprime la balise XML du bloc courant
Delete tree tags	: Supprime la balise courante et toutes celles qu'elle contient
Find ...	: Recherche d'une chaîne de caractères dans le texte
Preferences ...	: Configuration de l'éditeur

Style

New style ...	: Création d'une nouvelle feuille de style
Copy style ...	: Création d'une feuille de style en copiant la feuille courante
Load style ...	: Chargement en mémoire d'une feuille de style
Save current style ...	: Sauvegarde de la feuille de style courante
Remove current style	: Suppression en mémoire de la feuille de style
Edit styles ...	: Edition des caractéristiques de la feuille de style

Dtd

Load DTD ...	: Chargement d'une nouvelle DTD
Test DTD compatibility	: Validation partielle du document
Validate DTD	: Validation stricte du document

Il n'est pas a priori nécessaire que toutes les commandes du menu possèdent une icône, sauf éventuellement celles qui apparaissent également dans la barre d'icônes afin de permettre à l'utilisateur de les identifier plus facilement.

Barre icônes et état

La barre d'outils contient des commandes usuelles (qu'il faudra pouvoir faire évoluer facilement en fonction des besoins des utilisateurs), des informations générales (numéro de paragraphe), et la liste des feuilles de styles en mémoire ... Chaque groupe de fonctions doit être clairement séparé. Voici un exemple (non définitif) de barre d'outils :



Parmi les commandes nécessaires, on trouve :

- Les commandes courantes de manipulation de fichiers
- Les manipulations de style (édition)
- Les commandes d'édition de balise (extension/réduction de zones, delete tag ...)
- La sélection du style courant
- ...

Contenu de la fenêtre

Editeur de texte (TE)

OPTION : Il serait utile que l'éditeur de texte puisse se diviser horizontalement en deux (splitter), ainsi l'annotateur pourrait comparer deux parties non contiguës du document.

OPTION : Il serait utile de pouvoir créer plusieurs instances de l'éditeur (dans son ensemble) de manière à avoir plusieurs documents à l'écran sans qu'il soit nécessaire de lancer plusieurs fois l'application et donc plusieurs machines Java. Dans ce cas il faudrait rajouter une nouvelle commande dans le menu FILE : « New Editor ».

Visualisation XML (XM)

Les paramètres de visualisation du code XML (aspect, options par défaut ...) sont décrits dans la partie « configuration du système » de ce document.

Editeur d'attributs (EA)

Cette partie affiche le nom de la balise courante et la valeur des attributs associés. Les boutons de suppression d'étiquettes ne font plus partie de cette zone. La largeur des zones d'édition (menus et champs textuels) doit être limitée à 180 pixels environ.

OPTION : l'éditeur d'attributs et la liste des balises pourraient devenir de manière paramétrable deux fenêtres indépendantes redimensionnables. Une manière de procéder serait de rendre les parties EA et LB « détachables » de la fenêtre principale.

Listes des balises (LB)

Pour le contrôle du balisage, deux solutions sont possibles : soit désactiver les balises non autorisées (c'est la solution actuelle), soit rafraîchir directement la liste de manière à ne faire apparaître que les balises couramment utilisables. Cette seconde façon permet d'avoir une liste plus courte et lisible, mais pose des problèmes de réactivité de l'IU.

OPTION : Il faudrait pouvoir associer à chaque balise une « aide » qui s'afficherait classiquement lorsque le curseur reste quelques instants dessus. Les chaînes d'aides seraient décrites dans un fichier indépendant de la DTD (format à discuter).

Typologie des fichiers

L'éditeur fait ses sauvegardes (documents annotés, feuilles de styles, préférences, ...) dans des fichiers qui sont décrits à l'aide d'une syntaxe « XML ». Pour tous ces fichiers, le langage de description qui est utilisé (la DTD en d'autres termes) doit être explicitement nommé dans l'en-tête et doté d'un numéro de version :

```
<?xml version="1.0" ?>
<!DOCTYPE xanted-document-name PUBLIC "Version 1.0" "Url de la DTD">
```

Fichiers « Utilisateur »

L'éditeur est amené à manipuler les fichiers suivants :

- Des descriptions de langages d'annotations (DTD)
- Des feuilles de styles associées aux DTD (format XML)
- Les documents en cours d'annotation (format XML)
- Des fichiers de configuration (préférences) de deux sortes (format XML) :
 - Pour l'éditeur (chemin d'accès, polices, ...)
 - Pour chaque document annoté (DTD, feuilles de styles utilisées ...)

Les fichiers de configuration pour les documents annotés mémorisent la configuration de l'environnement de travail courant. Ainsi, lorsque l'utilisateur recharge un document il se retrouve dans la même configuration de travail. Ces informations ne sont pas sauvegardées dans le fichier annoté afin que celui-ci reste conforme à la DTD. Toutefois, la présence de ce fichier de configuration du document annoté est optionnelle afin de pouvoir travailler ; s'il est absent le fichier annoté est chargé dans l'environnement courant.

Fichiers « Système »

Le contenu du dossier contenant l'exécutable de l'éditeur est le suivant (aucun autre fichier n'a pas droit de figurer à ce niveau de la distribution ☺). Ce dossier doit contenir l'ensemble des données nécessaires à l'exécution.

- Xanted.jar : exécutable « double-cliquable »
- Xanted.bat : ligne de commande lançant l'exécutable
- /Preferences : fichiers par défaut

- Xanted.config : configuration de l'éditeur
- /Graphics/ : icônes et splash-screen
- /Documents/ : les documents de travail
 - DefaultDocument.xml
 - DefaultDocument.pref
 - ...
- /Styles/ : bibliothèque de styles
 - DefaultStyle.style
 - ...
- /DTD/ : bibliothèque de DTD
 - DefaultDTD.dtd
 - ...

Lors du lancement de l'application le document par défaut « DefaultDocument.xml » est chargé en mémoire³ ce qui provoque le chargement automatique de la DTD et des feuilles de style qui sont déclarées dans « DefaultDocument.pref ».

Gestion des « Feuilles de style »

Les feuilles de styles permettent d'associer aux balises de la DTD une apparence graphique (l'idée étant de reprendre ce qui est fait avec les CSS⁴) et une sémantique notamment pour ce qui concerne l'insertion automatique des balises. L'éditeur peut avoir plusieurs feuilles de styles en mémoire ce qui permet à l'utilisateur de passer de l'une à l'autre en fonction de la représentation qu'il souhaite avoir de son document.

Format des feuilles

On sauve chaque feuille dans un fichier spécifique en utilisant un format XML spécifique. Voici un exemple du format qui pourrait être utilisé pour chaque balise :

```
<?xml version="1.0" ?>
<!DOCTYPE xanted-style PUBLIC "Version 1.0" "Url de la DTD">
<style name ="nom du style">
...
  <balise nom="une_balise">
    <CSS code="le codage CSS de la balise"/>
    <Completion action ="" insert="" begin="" end="" />
  </balise>
...
</style>
```

Les attributs de la balise « Completion » correspondent aux options suivantes :

- action = nothing|complete
- insert = tag|section|document|expression
- begin = "expression régulière"
- end = "expression régulière"

³ Ce qui n'est pas le cas de la version actuelle ou il y a une DTD et une feuille de styles par défaut. En passant par un document, on obtient une configuration plus complète de l'éditeur en permettant notamment le chargement de plusieurs feuilles de style.

⁴ A plus long terme il faudrait sans doute regarder ce qui se fait autour de XSL.

Gestion des styles avec les documents

Lorsque l'on annote un document, les feuilles de styles courantes utilisées avec ce document doivent être sauveées avec lui. Cependant il est souhaitable de bien séparer l'annotation et les styles en d'associant à chaque fichier annoté un *fichier de préférences* contenant notamment la liste (sous la forme du chemin d'accès relatif) des fichiers de styles qu'il utilise et non la description des styles eux-mêmes. Ainsi il est possible de :

- de créer facilement des « bibliothèques de styles » indépendantes
- d'associer en outre des préférences « locales » à chaque fichier annoté

Il est à noter que lors de la sauvegarde d'un document annoté les feuilles de styles ne sont pas automatiquement sauveées (il faut qu'elles l'aient été explicitement soit via le menu « Edit » soit via la boîte de configuration des styles) et seules les feuilles de styles déjà sauvegardées sur disque sont mémorisées dans les préférences. Ainsi l'utilisateur peut facilement construire des styles qui n'ont d'existence que durant une session.

Configuration des documents

Le fichier de configuration d'un document est automatiquement créé lors de la sauvegarde d'un fichier annoté. Il contient les informations suivantes :

- Le chemin d'accès *relatif* de la DTD utilisée
- Les chemins d'accès *relatifs* des feuilles de styles en mémoire
- Les informations sur la position courante du curseur (TE et XM)
- Le style actif lors de la sauvegarde

Lors de la relecture d'un document annoté, ce fichier peut être manquant (ce peut être le cas si un utilisateur n'a distribué que le fichier annoté). Par contre en cas où il manque l'un des fichiers décrits dans la configuration, un message d'erreur est affiché.

Configuration du système

Le fichier « Xanted.config » permet à chaque utilisateur de définir son environnement de travail. On peut en distinguer les catégories de préférences suivantes :

- Le nom et l'emplacement des fichiers/dossiers par défaut
- Les polices par défaut
- Les couleurs par défaut
- Les diverses options pour paramétrer le comportement l'éditeur

Concernant les couleurs il n'est pas prioritaire de permettre une édition directe et on se contentera d'éditer le fichier de préférence « Xanted.config ». L'ordre dans lequel les couleurs sont déclarées dans ce fichier doit se retrouver dans l'interface

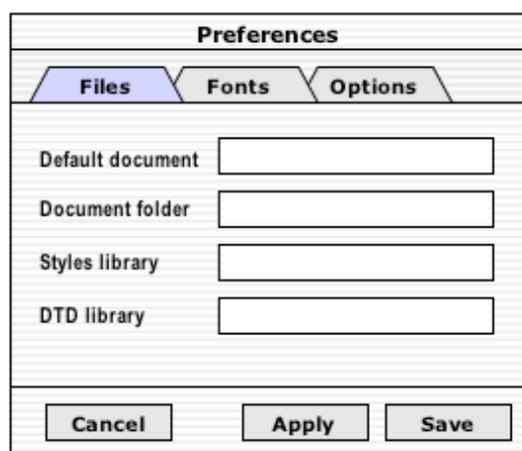
La boîte de dialogue « Préférences » (Menu EDIT) est non redimensionnable et structurée à l'aide « d'onglets » de manière à faciliter la tâche de l'utilisateur (et du programmeur). La boîte possède 3 boutons :

- « Cancel » : abandon des modifications en cours (on garde l'état antérieur)
- « Apply » : application des modifications en mémoire
- « Save » : application des modifications et sauvegarde du fichier de config

Onglet «Files» :

Il permet de définir les chemins d'accès (et/ou noms de fichiers) suivants :

- Document par défaut (le modèle) à charger lors du lancement de l'éditeur. Par défaut c'est le document : « /Preference/ DefaultDocument.xml »
- Le dossier de travail par défaut, c'est à dire l'emplacement où vont se ranger automatiquement les documents et leur préférences. Par défaut : « /Documents/ »
- L'emplacement d'une librairie de styles dans laquelle les feuilles de styles sont sauveées (menu STYLE). Par défaut : « /Styles/ »
- L'emplacement d'une librairie de DTD. Par défaut : « /DTDs/ »



Onglet «Fonts» :

Les préférences de styles de textes doivent porter sur les aspects suivants :

- Pour la fenêtre supérieure gauche (le texte)
 - Police par défaut du texte non annoté ("Plain text")
- Pour la fenêtre inférieure gauche (le balisage XML)
 - Police par défaut des balises ("Tag body")
 - Police par défaut des "fragments de phrases" ("Sample text")
- Initialisation des balises lors de la création d'un nouveau style vierge
 - Police par défaut des balises ("Default Tag Aspect")

Au niveau de la boîte de dialogue le plus simple est d'introduire un menu « pop-up » contenant les type de textes mentionnés ci-dessus avec en dessous le composant de sélection de police. C'est lui qui permet de sélectionner la police en cours de définition. Le sélecteur de police utilisé à utiliser ici doit être le même que celui qui est décrit pour les feuilles de styles pour sélectionner l'apparence d'une balise (gain de cohérence pour l'interface et gain en coût de développement).

Onglet «Options» :

Il permet de paramétrer les aspects suivants :

- Apparition des balises de fin dans la zone d'affichage du code XML. La checkbox est initialisée à « faux » (non cochée)
- Nombre maximum de caractères pour les balise affichée dans la zone XML (Tag Body). Par défaut 30 caractères.
- Nombre maximum de caractères pour les fragments de phrases dans la zone XML (Sample Text). Par défaut 50 caractères.
- Paramétrage du contrôle de l'insertion de balise dans le texte.
 - Strict : que des balises immédiatement plus spécifique (cas actuel)
 - Large : les balises plus spécifiques à un niveau quelconque de la DTD

Le menu FILE

- About : cette commande réaffiche l'écran d'identification (« splash-screen ») du logiciel. Il faut que *chaque nouvelle distribution* ait un numéro de version, ce numéro étant automatiquement surimposé en bas à droite de l'identification ...
- New document : création d'un nouveau document textuel vide. Le contexte de travail actuel (DTD et feuilles de styles) n'est pas modifié.
- Load text document : charge un fichier texte dans le contexte de travail courant.
- Load annotated document : recharge un document annoté dans l'éditeur et le contexte (DTD, Styles, ...) décrit dans le fichier de préférence s'il existe.
- Save annotated document : sauver le document annoté et son contexte.
- Export HTML document : sauvegarde du document sous la forme HTML. Le document sauvé est constitué de deux parties : 1) La déclaration des balises sous la forme de feuille de styles CSS (Class) 2) Le texte proprement dit annoté avec les types CSS. Voici un exemple du format attendu :

```

</head>
<style type="text/css" media="screen"><!--
.abstract { color: blue; font-style: italic; font-family: Arial; text-decoration: underline }
--></style>
</head>

<body bgcolor="#ffffff">
...
<p><span class="abstract">P2- In contrast, both the E213G and E213A Spo0A variants showed decreased
binding and completely abolished transcriptional activation of #spoIIa# and #spoIIe#, while the R214Gand
R214A variants completely abolished both DNA binding and transcriptional activation.
</span></p>
...
</body>

```

- Export RTF document : sauvegarde du document sous la forme RTF.
- Print : impression du document courant
- Quit : Fermeture du logiciel avec apparition d'une boîte de dialogue non redimensionnable demandant si l'on veut sauvegarder le document. Il y a trois boutons : « Cancel », « Quit » et « Save and Quit ». Cette boîte ne doit cependant apparaître que si le document a été modifié depuis la dernière sauvegarde.

Le menu EDIT

- Gestion du copier/coller

De par l'utilisation qui doit être faite de l'éditeur d'annotation il est nécessaire de pouvoir effectuer des copier/coller de blocs contenant éventuellement des balises. Toutefois, on ne doit pouvoir coller du texte « balisé » qu'à des endroits compatibles avec la DTD courante. Afin de simplifier le problème on peut implémenter deux types de commandes : l'une pour manipuler du texte « brut » et l'autre pour manipuler du texte balisé.

Dans le premier cas le couper/copier/coller peut être effectué librement : lors du collage, le texte est simplement rajouté (sans ses balises initiales) à l'endroit de l'insertion. L'emplacement de l'insertion est choisi soit en cliquant dans l'éditeur, soit lorsque plusieurs balises commencent (ou finissent) au même endroit en sélectionnant dans la zone XML le nom de la balise où l'on souhaite faire l'insertion : la première sélection place le curseur en début de balise, la seconde en fin (et ainsi de suite en boucle)

Dans le second cas « avec balises » les commandes de couper-copier agissent de la manière suivante : le bloc à copier est soit celui dans lequel le curseur est placé (en l'absence de toute sélection explicite), soit la sélection de bloc courante (qui peut être activé via un double clic dans la zone XML). Si ce n'est pas le cas la commande est ignorée avec une boîte d'alerte expliquant la raison de cet échec.

Lors du collage, la balise englobant le bloc sélectionné doit être compatible (c'est à dire plus « spécifique » au sens large) que celle correspondant au point d'insertion.

- Delete current tag : suppression de la balise courante, c'est à dire celle contenant le caractère d'insertion ou le bloc couramment sélectionné.
- Delete tree tags : suppression de la balise courante et, récursivement, de toutes les balises qu'elle contient. Comme cette commande est potentiellement dangereuse (risque de perdre l'ensemble des annotations), il serait raisonnable de la faire confirmer par une simple boîte de dialogue « OK – Cancel », à moins que la touche SHIFT ne soit pressée (auquel cas la commande est directement exécutée).
- Find : Appel d'une BD permettant à l'utilisateur de faire une recherche de chaîne de caractères. Il serait souhaitable de prévoir la possibilité de faire une recherche de l'occurrence précédente (bouton « previous ») et suivante (bouton « next »)
- Préférences : appel la boîte de configuration de l'éditeur (cf. description)

Gestion du menu «Style

Le style courant, c'est-à-dire le style qui est *effectivement utilisé* pour colorer les balises, doit être signalé (par exemple, l'étiquette ou le corps du bouton actif est colorée différemment – à voir en fonction des possibilités en JAVA) dans la barre des icônes. Pour passer d'un style à l'autre il suffit de cliquer sur le bouton correspondant

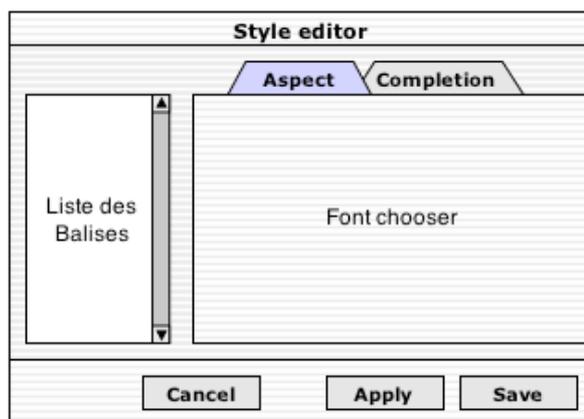
- New style : création d'un nouveau style et ajout d'un nouveau bouton dans la barre des icônes. Chaque nouveau style doit avoir un nom qui est utilisé pour étiqueter le

bouton correspondant dans la barre des icônes. Le style des balises est initialisé avec la police qui est décrite par l'attribut « Tag default Aspect » dans les préférences.

- Copy style : création d'un nouveau style en copiant sur le style courant (actif) et ajout d'un nouveau bouton dans la barre des icônes.
- Load style charge un fichier de style précédemment sauvé et ajoute automatiquement un nouveau bouton dans la barre des icônes (à la fin) dont l'étiquette est le nom du fichier (sans l'extension du fichier bien évidemment). Si le nom existe déjà dans la barre de style le nouveau style doit remplacer l'ancien. Noter qu'une feuille de style ne contient pas forcément toutes les balises (éventuellement aucune) de la DTD courante, dans ce cas on n'initialise que les balises en communs, et la partie description sémantique (insertion automatique) n'est pas prise en compte.
- Save current style Sauve le style couramment utilisé (via la BD de sauvegarde « standard »). Si le fichier existe déjà on le remplace par le nouveau, sinon on crée un nouveau fichier. Les styles sont sauvés soit par cette commande soit lors de l'édition.
- Delete current style : Supprime le style courant (mais pas le fichier), on utilise alors le style précédent dans la liste on doit toujours laisser au moins un style dans la barre (par exemple le dernier style n'est pas supprimable) ...
- Edit Style : Appelle la boîte d'édition de style décrites ci-dessous

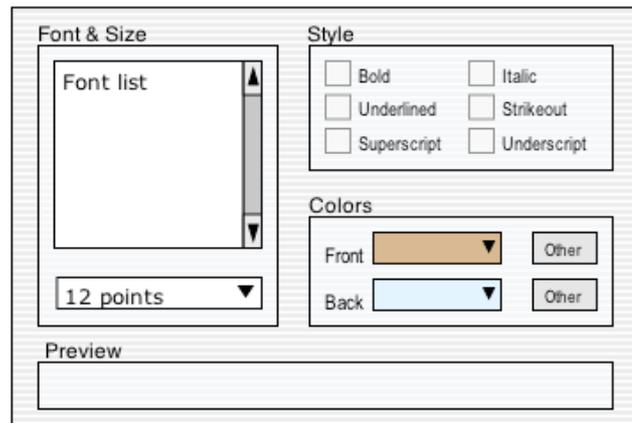
On utilisera une boîte de dialogue non redimensionnable pour modifier les paramètres graphiques et sémantiques associés à chacune de balises de la DTD. A gauche, on affiche la liste des balises de la DTD courante. Afin de simplifier la tâche de l'utilisateur, les balises doivent apparaître *dans le même ordre* que dans la fenêtre principale. L'onglet « Aspect » contient un sélecteur de police et l'onglet « Completion » les options de configuration de l'ajout automatique de styles. La boîte possède 3 boutons :

- « Cancel » : abandon des modifications
- « Apply » : application des modifications en mémoire
- « Save » : application des modifications et sauvegarde du fichier de style



Sélection de l'aspect du style

A chaque balise on peut faire correspondre une police et une taille, un aspect (gras, italique, ...) et une couleur de caractère et de fond. Voici la BD de sélection :



Dans le choix des couleurs, le bouton « Other » à droite de chaque menu (front et back) donne accès à un « Color Chooser⁵ » permettant de s'affranchir des couleurs prédéfinies.

Complétion automatique

Dans la DTD utilisée dans Caderige (mais ce n'est pas spécifique), le marquage d'un agent d'interaction, d'une cible, d'une interaction ... est relativement naturel à effectuer pour l'utilisateur (biologiste) puisqu'il consiste à assigner un sens biologique aux éléments du texte. Par contre, l'introduction de certaines balises XML plus générales (par exemple ABSTRACT, SENTENCE, ...) est moins intuitive et risque d'être assez systématiquement oubliée ce qui conduira à devoir effectuer de nombreuses corrections lors de la vérification. Par ailleurs, on constate que certaines balises sont sémantiquement « couplées » dans la DTD : par exemple, une balise d'agent générique d'interaction <GA> est forcément englobée dans un fragment <GAF>.

La complétion automatique permet de résoudre ces différents problèmes, en associant aux feuilles de style des directives (optionnelles) d'insérant automatiquement les balises de niveau supérieur (englobantes) lorsqu'elles sont absentes dans le document. Cette information n'est « qu'en partie » déductible de la DTD, car si celle-ci permet de connaître les balises englobantes, elle ne permet pas de savoir OU elles doivent être introduites dans le document. Voici les options que l'on peut associer à chaque balise.

Si la balise de niveau immédiatement supérieur est manquante :

- Ne rien faire
- Insérer [automatiquement|interactivement]cette balise
- Autour de la balise courante
- Autour du paragraphe courant
- En début et fin du document
- En utilisant les expressions régulières suivantes
 - Recherche du début :
 - Insertion [Avant|Après]
 - Recherche de la fin :
 - Insertion [Avant|Après]

On trouve deux séries d'options imbriquées (radio-boutons) dans la boîte de dialogue. Tout d'abord pour décider si l'absence de l'étiquette de niveau supérieur déclenche une

⁵ J'imagine qu'il doit y en avoir un composant standard dans les bibliothèques JAVA ...

action ou non, puis, si c'est le cas, pour indiquer où sera insérée la balise supérieure manquante. Via un pop-up l'utilisateur décide si l'insertion est automatique ou avec une demande de validation explicite via une BD.

La dernière option permet d'effectuer une recherche plus fine des points d'insertion à l'aide d'expressions régulières. La recherche du point d'insertion de la balise ouvrante s'effectue en recherchant vers le début du document à partir de la gauche de la balise ouvrante actuelle (et inversement pour la balise fermante). Via deux menus pop-up on sélectionne l'endroit de l'insertion vis à vis des expressions trouvées. Si l'une des expressions n'est pas trouvée rien ne se passe (on ne pas doit insérer une seule balise).

Ainsi, l'exemple ci-dessus est la configuration de la feuille de style associée à une balise <GAn> de manière à ce qu'on ajoute interactivement (c'est à dire avec une confirmation de l'utilisateur via une BD) ou automatiquement (sans confirmation)⁶ la balise <GAfn> qui doit l'englober autour de la balise <GAn> introduite en premier.

En pratique, cette gestion de l'insertion des balises manquantes doit être effectuée de manière récursive jusqu'à ce qu'on arrive à la racine de la DTD ou jusqu'à ce que l'on retombe sur une balise qui était effectivement prévue par la DTD. Les attributs des balises automatiquement insérés, prennent les valeurs par défaut. Par exemple, si l'utilisateur charge un document et commence directement à introduire une balise <GA1> dans une phrase quelconque, cela pourrait déclencher les insertions suivantes en cascade (pour peu évidemment que les feuilles de styles aient été correctement configurées) :

- Ajout de la balise <GAF1> autour de <GA1>
- Ajout de la balise <GENIC-INTERACTION> autour du paragraphe courant
- Ajout de la balise englobante <ANNOTATED-DOCUMENT>

Gestion du menu « Dtd »

- Load DTD : charge une DTD en mémoire. Cette opération réinitialise l'environnement de travail (suppression des styles et des balises).
- Test DTD compatibility : effectue une vérification « faible » de la DTD. C'est à dire que l'ordre d'imbrication des balises est correcte vis à vis de la DTD, indépendamment du fait que « l'ensemble » du document est conforme ou non.
- Validate DTD : vérifie que le document est strictement conforme à la DTD. Chaque problème rencontré doit être clairement signalé : nom de la balise erronée (ou manquante) et numéro de ligne (paragraphe correspondant). La boîte d'erreur doit permettre (via un bouton) à l'utilisateur de se positionner directement sur l'endroit problématique afin de le corriger.

⁶ Le choix entre les deux modes s'effectue via un pop-up menu.