

Le compilateur NAZE

Alexandre LHUILLIER

FIPA 2

21 juin 2011

Table des matières

1	Présentation du sujet	1
1.1	But	2
2	Manuel d'utilisation	2
2.1	Fichier d'entrée	2
2.2	Champs obligatoires	3
2.3	Champs communs	3
2.4	Champs spécifiques aux disques	4
2.5	Champs spécifiques aux livres	4
2.6	Commentaires	5
2.7	Messages d'erreurs et solutions correspondantes	5
2.8	Fichier de sortie	7
3	Conclusion	9
3.1	Fonctionnalités présentes	9
3.2	Améliorations possibles	10
A	Exemple complet de fichier d'entrée	11
B	Instructions de compilation	13

1 Présentation du sujet

L'application est un compilateur qui, à partir d'un fichier texte - dont le format sera décrit dans le manuel d'utilisation - , produit plusieurs fichiers HTML représentant une médiathèque composée de descriptions de livres et de disques.

1.1 But

Le but du projet derrière cette application est de se familiariser avec les applications pratiques de la théorie des langages, en créant une grammaire, un lexique et à partir de là, un compilateur.

2 Manuel d'utilisation

Le compilateur NAZE s'exécute avec une commande de ce format :

```
./naze <fichier d'entrée> <répertoire de sortie>
```

Si le répertoire de sortie n'existe pas ou si l'utilisateur n'a pas les droits en écriture dessus, le programme abandonne.

Le format du fichier d'entrée est décrit dans la suite.

2.1 Fichier d'entrée

Le fichier d'entrée est un fichier texte composés de plusieurs blocs délimités par des parenthèses. Ils ont pour titre "LIVRE" ou "DISQUE" selon le type de média qui y est décrit. Ce titre est seul sur une ligne.

Un bloc est ensuite composé de plusieurs champs, de la forme

```
\<nom_balise> = <valeur_champ>
```

Il peut y avoir n'importe quel nombre de blancs quelconques entre le nom de la balise, le "=" et la valeur.

```
(  
DISQUE  
.  
.  
.  
)  
(  
LIVRE  
.  
.  
.  
)
```

Extrait d'un fichier d'entrée composé de deux blocs.

2.2 Champs obligatoires

Dans chaque bloc, un champ `\Titre` et un champ `\Auteurs` sont obligatoires et ne doivent être présents qu'une fois chacun. Leur ordre et leur emplacement dans le bloc n'a pas d'importance.

```
(
DISQUE
\Auteurs=White Jack ; White Meg
\Titre=Icky Thump
.
)
(
LIVRE
.
\Titre = Trois vies chinoises
.
\Auteurs =Dai Sijie
)
```

Le même fichier avec les balises `\Titre` et `\Auteurs`, placées à des emplacements quelconques.

La valeur d'un champ `\Titre` est un texte libre.

La valeur du champ `\Auteurs` est :

Pour un auteur, de la forme `<1 nom de famille> <0 ou 1 ou plusieurs prénoms>`

Les prénoms peuvent être séparés par un blanc quelconque (espace(s), tabulation(s) ou saut(s) de ligne).

Pour plusieurs auteurs : de la forme `<auteur> ;<auteur> ;...;<auteur>`

2.3 Champs communs

Il existe plusieurs champs communs aux disques et aux livres.

Voici les type de données qu'ils peuvent accepter :

- ENTIER : un nombre entier
- URL : une chaîne de caractères qui représente une URL. Si elle contient des blancs, ils seront supprimés lors de la compilation.
- TEXTE LIBRE : Une chaîne de caractères libre, la plupart des caractères sont acceptés. Il est tout de même préférable d'éviter les accents. Pour écrire une parenthèse sans provoquer d'erreur de syntaxe, elle doit être préfixée d'un " `\`" (notez l'espace avant le backslash. Un backslash

seul était prévu au départ, mais pas moyen de faire accepter un fichier sans l'espace actuellement.)

Voici les balises communes, en plus des 2 balises obligatoires, et les types de données qu'elles acceptent :

- \Annee : ENTIER
- \Photo : URL
- \Notes : TEXTE LIBRE

2.4 Champs spécifiques aux disques

Deux types de données supplémentaires existent pour les disques :

- DUREE : <entier> :<entier>
- INTERPRETES : Pour un interprète seul : <nom> <prenoms>(qualité)
Pour plusieurs interprètes, ils sont séparés par des points-virgules. La qualité est un TEXTE LIBRE.

Voici les champs et leur type de donnée correspondant :

- \Interpretes : ENTIER
- \Plage : TEXTE LIBRE
- \Duree : DUREE

Une Plage peut être suivie ou non d'une durée. Un champ Durée ne peut pas être présent sans champ Plage le précédant. Plusieurs champs Plage (chacun éventuellement suivi d'une durée) peuvent être présent dans un bloc DISQUE.

Exemple de bloc DISQUE :

```
(  
DISQUE  
\Auteurs=White Jack; White Meg  
\Titre=Icky Thump  
\Interpretes=White Jack(guitare et chant); White Meg(batterie)  
\Plage = Catch Hell Blues  
\Duree = 3 :00  
\Plage = Little Cream Soda  
)
```

2.5 Champs spécifiques aux livres

Voici les champs et leur type de donnée correspondant :

- \Resume : TEXTE LIBRE

– \Editeur : TEXTE LIBRE

Exemple de bloc LIVRE :

```
(
LIVRE
\Editeur= Flammarion
\Titre = Par une nuit ou la lune ne s'est pas levee
\Resume = Une histoire de langues inconnues, de chasse au tresor,
de trucs sympas.
\Auteurs =Dai Sijie
)
```

2.6 Commentaires

Vous pouvez inclure des commentaires dans le fichier d'entrée. Pour cela, vous avez deux méthodes :

– Pour entrer un commentaire qui s'étend jusqu'à la fin d'une ligne, commencez votre commentaire par "%%".

Exemple :

```
%%Tout sera commenté jusqu'au prochain retour à la ligne
```

– Pour entrer un commentaire qui peut s'étendre sur plusieurs lignes, encadrez-le avec "(" au début et ")" à la fin.

Exemple :

```
(* Dans cette ville,
La compil',
c'est subtil,
pas facile. *)
```

2.7 Messages d'erreurs et solutions correspondantes

Le programme abandonne la compilation dès la première erreur rencontrée.

Chaque erreur est précédée d'un message :

```
erreur ligne <numero de ligne> : <dernier mot détecté>
```

Voici les messages qui peuvent suivre :

Le champ obligatoire \Auteurs est déjà présent plus haut dans ce media

Un seul champ "\Auteurs" doit être présent dans un bloc. Ce message signifie que vous en avez plusieurs dans un bloc.

Le champ obligatoire \Titre est déjà présent plus haut dans ce media
Un seul champ "\Titre" doit être présent dans un bloc. Ce message signifie que vous en avez plusieurs dans un bloc.

Le champ obligatoire \Auteurs est absent de ce media
Le champ \Auteurs est obligatoire dans un bloc. Ce message signifie que le bloc à la ligne signalée n'en possède pas.

Le champ obligatoire \Titre est absent de ce media
Le champ \Titre est obligatoire dans un bloc. Ce message signifie que le bloc à la ligne signalée n'en possède pas.

Un champ \Resume ne doit pas être présent dans un bloc DISQUE
Un champ \Editeur ne doit pas être présent dans un bloc DISQUE
Un champ \Plage ne doit pas être présent dans un bloc LIVRE
Un champ \Duree ne doit pas être présent dans un bloc LIVRE
Un champ \Interpretes ne doit pas être présent dans un bloc LIVRE
Ce type de message apparaît lorsqu'un champ spécifique à un type de bloc apparaît dans un bloc de l'autre type.

Un champ \Duree doit être précédé d'un champ \Plage
Un champ \Duree est présent sans \Plage le précédant à la ligne indiquée.

champ precedent mal forme
format du champ <nom du champ> : <nom du champ> = <type du champ>
Ce message signifie qu'un champ mal formé a été trouvé. Selon le cas, <nom du champ> est remplacé par le nom de la balise (\Titre, ...) et <type du champ> par le type qui lui est associé (<Texte libre>, <nombre entier>...).

2.8 Fichier de sortie

Le premier fichier généré est le fichier index.html, contenant un sommaire de toutes les oeuvres :

Votre Mediatheque

Doctor Who OST Series 1 & 2	M. Gold	3 minutes 2 secondes
Icky Thump	M. White et J. White	36 minutes 49 secondes
Le Complexe de Di	S. Dai	
Par une nuit ou la lune ne s'est pas levée	S. Dai	Gallimard

Les oeuvres sont triées par ordre alphabétique de leur nom, sans prise en compte du type de média. Les auteurs apparaissent avec les initiales de leurs prénoms et leur nom de famille complet. Ils sont triés par ordre alphabétique de leurs noms et prénoms. La troisième colonne contient, s'ils sont renseignés, soit l'éditeur si le média est un livre, soit la durée totale d'écoute, calculée avec la somme des durées des plages, si c'est un disque

A chaque oeuvre correspond une page détaillée comportant plus d'informations sur l'oeuvre.

Si le champ \Photo est rempli dans le fichier d'entrée, l'image qui est à l'URL renseignée est affichée sur la page. Notez que le programme ne vérifie pas si le fichier référencé par l'URL est bien une image.

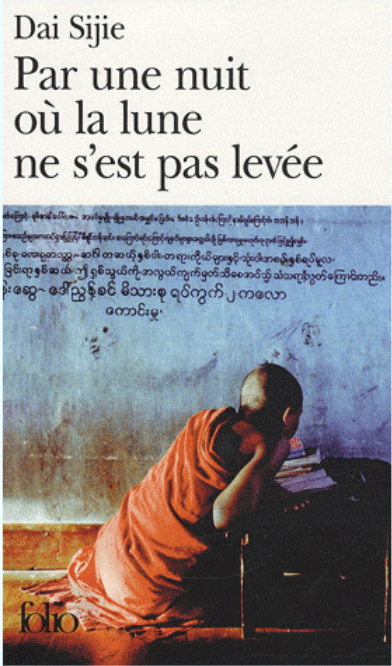
Un exemple de page détaillée sur un livre :

Par une nuit ou la lune ne s'est pas levée

S. Dai

Gallimard

Dai Sijie
Par une nuit
où la lune
ne s'est pas levée



Une histoire de chasse au trésor, avec des chinois, des français ...

Un exemple de page détaillée sur un disque :

M. White et J. White

Icky Thump
2007
M. White (Batterie) et J. White (Guitare chant)



1 - Little Cream Soda [3.45]
2 - Conquest [2.48]
3 - Rag And Bone [3.48]
4 - Effect And Cause [3.00]
5 - I'm Slowly Turning Into You [4.34]
6 - St Andrew (This Battle Is In The Air) [1.49]
7 - Prickly Thorn, but Sweetly Worn [3.05]
8 - Catch Hell Blues [4.18]
9 - 300 MPH Torrential Outpour Blues [5.28]
10 - Icky Thump [4.14]

mes notes

3 Conclusion

3.1 Fonctionnalités présentes

L'application est actuellement capable de produire plusieurs fichiers .html à partir d'un fichier d'entrée : un fichier index, avec la liste des médias, et un fichier avec une description détaillée de chaque oeuvre.

L'utilisateur peut insérer des commentaires dans son fichier d'entrée pour apporter des précisions qui ne doivent pas être visibles en sortie.

Les messages d'erreur qui apparaissent à la compilation permettent de facilement localiser le problème dans le fichier d'entrée. Lorsque le message d'erreur ne propose pas de solution, il communique au minimum la ligne de l'erreur rencontrée et le dernier mot détecté.

3.2 Améliorations possibles

Actuellement, le code HTML est présent en dur dans le programme. On pourrait imaginer ajouter, en option du programme, un fichier HTML qui servirait de template pour pouvoir personnaliser la mise en page de la médiathèque.

On pourrait également mettre plusieurs fichiers d'entrée pour générer en sortie une seule collection de fichiers contenant les description des médias de tous les fichiers d'entrée.

Une autre amélioration possible serait un projet en elle-même : il s'agirait de rendre le format du fichier d'entrée extensible.

Concrètement, on aurait d'un côté un fichier listant tous les types de médias possibles, toutes les balises possibles avec leur type et des informations complémentaires :

```
\Titre TEXTELIBRE OBLIGATOIRE ALL
```

```
\Annee ENTIER FACULTATIF ALL
```

```
\Plage TEXTELIBRE DISQUE
```

```
.
```

```
.
```

On aurait, de l'autre, un fichier template qui dirait où chaque champ est placé dans le HTML.

Un programme générerait à partir de ceux-ci les fichiers lex et yacc correspondants, et des fichiers .c pour la gestion de champs plus complexes (comme les auteurs ou les plages, qui peuvent être présents plusieurs fois) et pour l'écriture du HTML.

Les fichiers générés seraient directement compilables pour obtenir un nouveau compilateur.

A Exemple complet de fichier d'entrée

```
vraiexemple.txt :
(
DISQUE%%commentaire a la fin du libelle
\Auteurs = White Jack;White Meg
\Interpretes = White Jack(Guitare chant);White Meg(Batterie)
%%commentaire inutile
\Titre=Icky Thump
\Plage=Icky Thump
\Duree = 4 :14
\Plage=300 MPH Torrential Outpour Blues
\Duree=05 :28
\Plage=Catch Hell Blues
\Duree=04 :18
\Plage=Prickly Thorn, but Sweetly Worn
\Duree=3 :05
\Plage=St Andrew \ (This Battle Is In The Air \)%%Bordel, un vrai titre avec
des parentheses dedans !
\Duree=01 :49
\Plage=I'm Slowly Turning Into You
\Duree=04 :34
\Plage=Effect And Cause
\Duree=03 :00
\Plage = Rag And Bone
\Duree = 3 :48
\Plage =Conquest
\Duree = 2 :48
\Plage = Little Cream Soda
\Duree=0003 :045
\Notes=mes notes
\Photo = http ://www.urb.com/uploads/reviews/cd_reviews/The_White_Stripes_Icky_Thump_Third_Man
_Warner_Bros..jpg
\Annee = 2007
)
(
(*)
un commentaire
sur plusieurs lignes
*)
LIVRE
```

```

\Editeur=Hachette
\Photo=http://www.decitre.fr/gi/64/9782070358564FS.gif
%%espaces, tabulations et sauts de ligne supprimes durant la lecture d'une url
\Auteurs=Dai Sijie %%Jolie balise \Auteurs
\Editeur =Gallimard%%Cet editeur ecrase le premier
\Resume=Un livre tres interessant
\Titre = Par une nuit ou la lune ne s'est pas levee(*la comme j'ecris un 2eme
resume
le 1er sera ecrase*)
\Resume=Une histoire (*commentaire en plein milieu du resume 2*2*)
de chasse au tresor, avec des chinois, des francais ...)
(%%le 3eme media
DISQUE
\Interpretes=BBC \Auteurs=Gold Murray
\Titre= Doctor Who OST Series 1 & 2
\Annee =2007
\Plage= Doomsday
\Duree=3 :2
\Plage = Seeking the Doctor
\Plage = I'm Coming To Get You
\Plage = Rose's Theme
\Plage=Father's Day
\Plage = Madame de Pompadour
\Photo=http://2.bp.blogspot.com/_LfbRKURMnyM/TSUNf6yBI1I/AAAAAAAAAug/h5zjimZj8ss/s1600/
Doctor+Who+Series+1-2+front+cover.jpg
)(
LIVRE
\Titre = Le Complexe de Di
\Auteurs = Dai Sijie
\Resume = Un juge chelou qui accepte de droles de corruptions, un chinois qui
a vécu en France qui veut libérer une détenue, et une scène qui lui a fait mériter
son prix Femina.
)

```

B Instructions de compilation

Commandes à effectuer dans le dossier PICO_LHULLIER.

Recompiler l'exécutable :

```
make default
```

Supprimer les fichiers intermédiaires :

```
make clean
```

Supprimer l'exécutable :

```
make delete
```