

S'y retrouver dans la jungle des Computer Modern

Eddie Saudrais

Le 17 septembre 2005

Codages T1 et OT1

Le codage, c'est l'organisation des caractères d'une fonte dans un tableau, chaque caractère étant repéré par son numéro. Le premier codage disponible avec T_EX est le codage OT1, sur 7 bits : seuls 128 caractères sont disponibles, ce qui ne permet pas d'y inclure les caractères accentués.

```
Γ Δ Θ Λ Ξ Π Σ Υ Φ Ψ Ω ff fi fl ffi ffi
1 j ` ´ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘
- ! " # $ % & ' ( ) * + , - . /
0 1 2 3 4 5 6 7 8 9 : ; j = i ?
@ A B C D E F G H I J K L M N O
P Q R S T U V W X Y Z [ " ] ^ `
‘ a b c d e f g h i j k l m n o
p q r s t u v w x y z - — “ ~ ”
```

TAB. 1 – Le codage OT1

Les caractères accentués sont alors construits, à l'aide de la macro `\accent` par l'assemblage des deux caractères : la lettre et l'accent. Cette méthode ne permet pas la césure des mots accentués.

Le codage T1, sur 8 bits, contient 256 caractères, ce qui permet d'y inclure les caractères accentués.

```
` ´ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘ ˘
“ ” „ « » - — o i j ff fi fl ffi ffi
˘ ! " # $ % & ' ( ) * + , - . /
0 1 2 3 4 5 6 7 8 9 : ; < = > ?
@ A B C D E F G H I J K L M N O
P Q R S T U V W X Y Z [ \ ] ^ _
‘ a b c d e f g h i j k l m n o
p q r s t u v w x y z { | } ~ -
Ǻ Ȧ Ć Č Ď Ě Ę Ğ Ĺ Ł ǻ Ń Ņ Ő Œ Ó Ŕ
Ř Ś Š Ÿ Ŧ Ŧ Ũ Ū Ÿ Ž Ž Ž Œ Œ Œ Œ
ǻ ą ę ę ę ę ğ Í Ĭ ĭ ů ů ů ů ů ů ů ů
ř ś š š ŧ ŧ ů ů ŷ ž ž ž ij j j £
À Á Â Ã Ä Å Æ Ç È É Ê Ë Ì Í Î Ï
Ð Ñ Ò Ó Ô Õ Ö Ø Ù Ú Û Ü Ý Þ ß
à á â ã ä å æ ç è é ê ë ì í î ï
ð ñ ò ó ô õ ö ø ù ú û ü ý þ ß
```

TAB. 2 – Le codage T1

C'est le codage qu'il faut utiliser pour composer des langues européennes (comme le français).

Il existe bien sûr d'autres codages (TS1 pour accéder à des symboles supplémentaires, des codages propres à d'autres alphabets, de codages mathématiques).

Pour composer du texte courant, on sélectionne le codage utilisé en écrivant `\usepackage[OT1]{fontenc}` ou `\usepackage[T1]{fontenc}` dans le préambule. Si l'on ne précise rien, c'est le codage OT1 qui est sélectionné par défaut.

Les fontes commerciales sont livrées dans un codage propre à Adobe (8r). Lorsqu'on les adapte à T_EX, on peut construire une fonte au codage OT1 ou une fonte au codage T1. À l'heure actuelle, il n'y a pas de raison de se limiter au codage OT1 pour écrire en français.

Métafontes et fontes Postscript

T_EX permet d'utiliser deux types de fontes (je laisse de côté les fontes TrueType prises en charge par pdfL^AT_EX) : les métafontes et les fontes TrueType.

Les métafontes sont des fontes créées avec le programme METAFONT de Knuth. Bien que la description des caractères soit vectorielle (elle est faite à l'aide de courbes de Bézier), DVIPS va les numériser (fichiers pk) lors de la création du fichier ps. Cela ne pose pas de problème pour une utilisation du fichier ps, mais **un fichier pdf obtenu en utilisant une métafonte sera très moche**¹.

Les fontes postscript sont vectorielles, et ne sont numérisées à aucun stade de la création de fichier pps ou pdf. **Il faut utiliser des fontes postscript de type 1 pour obtenir des fichiers pdf de bonne qualité.**

Attention : T_EX ne voit pas la différence entre une métafonte et une fonte postscript. Il ne travaille pas avec les dessins des caractères, mais uniquement avec les dimensions des boîtes qui les contiennent (c'est-à-dire avec les fichiers de métrique tfm). C'est DVIPS qui fera la distinction.

Le cas particulier de la famille Computer Modern

Computer Modern au codage OT1

La famille Computer Modern a été dessinée par Knuth en s'inspirant de la fonte Monotype Modern 8A, à l'aide de son programme METAFONT.

Cette famille Computer Modern n'existe qu'au codage OT1. Une conversion au format postscript de type 1 est disponible librement.

Si vous ne mettez rien dans le préambule : la famille Computer Modern sera utilisée, au codage OT1. Que vous ayez installé les fichiers pfb (dessin postscript des caractères) ou non ne change rien pour la compilation du document : les fichiers de métrique sont les mêmes (cmr12.tfm en corps 12 par exemple).

Quand vous appelez ensuite DVIPS :

- si vous n'avez pas installé les fichiers pfb, et signalé leur existence à DVIPS (fichier psfonts.map), les fichiers bitmaps pk seront utilisés ;
- si les fichiers pfb sont installés correctement, DVIPS va inclure la description postscript des caractères dans le fichier ps.

La version postscript de type 1 de la famille Computer Modern est mise à disposition par Bluesky. Dans une distribution correcte, on trouvera les fichiers pfb dans `textmf/fonts/type1/bluesky/cm`.

Computer Modern au codage T1 : European Computer

Une adaptation de la famille Computer Modern au codage 8 bit a été réalisé : c'est la famille EC (European Computer). Il ne s'agit pas d'un simple recodage : les dessins de certains caractères et des accents sont légèrement différents, comme on peut le voir ci-après.

¹Avec une distribution récente, je trouve que le résultat tend à devenir acceptable, quand il ne s'agit que de quelques caractères dans un document.

CM	À	É	Ê	ß
EC	À	É	Ê	ß

Les métriques sont sensiblement différentes :

CM	abcdefghijklmnopqrstuvwxy
EC	abcdefghijklmnopqrstuvwxy

Ces fontes sont longtemps restées disponibles au format métafont uniquement ; il n’y avait pas de solution libre pour utiliser de version Postscript des Computer Modern au codage T1.

Les choses changent ; on a deux possibilités pour utiliser la famille Computer Modern en codage T1 au format postscript :

- une conversion des fontes EC au format postscript est librement distribuée : ils s’agit des fontes cm-super. Elles ne sont pas d’excellente qualité² (elles sont été obtenues par des logiciels de conversion) ;
- une famille de substitution de qualité a été développée : il s’agit de lmodern. Dans l’état actuel des choses, la famille lmodern est moins complètes que l’ensemble EC, mais elle est en cours de développement.

CM (postscript (Bluesky), codage OT1)	À	É	Ê	ß
lmodern (CM postscript, codage T1)	À	É	Ê	ß
EC (métafonte, codage T1)	À	É	Ê	ß
cm-super (EC postscript, codage T1)	À	É	Ê	ß

Pour utiliser lmodern, il faut appeler le package correspondant : `\usepackage{lmodern}`.

Résumons

Codage ou extension sélectionné	Fonte utilisée
<code>\usepackage[OT1]{fontenc}</code>	Computer Modern 7 bits : — postscript si installée (Bluesky) — CM (métafonte) sinon
<code>\usepackage[T1]{fontenc}</code>	European Computer 8 bits : — cm-super si installée (conversion de EC en postscript) — EC (métafonte) sinon
<code>\usepackage{lmodern}</code>	lmodern (postscript, dessin des CM au codage T1)

- La famille lmodern n’est disponible qu’au codage T1 ; le chargement du package lmodern active automatiquement ce codage.

Et ce document, il est composé en quoi ?

Le texte courant est composé en Utopia, à l’aide de l’extension fourier (fourier est une famille complète, à base d’utopia, comprenant tous les symboles mathématiques utiles).

La fonte sans empattement utilisée pour les titres est officina sans. C’est aussi la fonte utilisée pour les noms de commandes ou d’extension de fichiers — on compose tfm en écrivant `\texttt{tfm}` —, réduite à 95 % de sa taille pour obtenir la même hauteur d’œil qu’Utopia.

²Il faut vraiment être difficile pour ne pas s’en contenter !