

xstring

v1.81

Manuel de l'utilisateur

Christian TELLECHEA
unbonpetit@netc.fr

9 décembre 2018

Résumé

Cette extension, qui requiert Plain ε -TeX, regroupe un ensemble de macros manipulant des chaînes de tokens (ou lexèmes en français). Les macros peuvent être utilisées de façon basique dans le traitement des chaînes de caractères mais peuvent également être utiles en programmation TeX pour la manipulation des tokens, c'est-à-dire du code TeX. Parmi les fonctionnalités, les principales sont :

- ▷ des tests :
 - une chaîne en contient elle une autre au moins n fois ?
 - une chaîne commence t-elle ou finit-elle par une autre ? etc.
 - une chaîne représente t-elle un entier relatif ? Un nombre décimal ?
 - deux chaînes sont-elles égales ?
- ▷ des recherches de chaînes :
 - recherche de ce qui se trouve avant (ou après) la n^{e} occurrence d'une sous-chaîne ;
 - recherche de ce qui se trouve entre les occurrences de 2 sous-chaînes ;
 - sous-chaîne comprise entre 2 positions ;
 - recherche d'un groupe entre accolades par son identifiant.
- ▷ le remplacement de toutes ou des n premières occurrences d'une sous-chaîne par une autre sous-chaîne ;
- ▷ des calculs de nombres :
 - longueur d'une chaîne ;
 - position de la n^{e} occurrence d'une sous-chaîne ;
 - comptage du nombre d'occurrences d'une sous-chaîne dans une autre ;
 - position de la 1^{re} différence entre 2 chaînes ;
 - renvoi de l'identifiant du groupe dans lequel une recherche ou une coupure s'est faite.

D'autres commandes permettent de gérer les tokens spéciaux normalement interdits dans les chaînes (# et %), ainsi que d'éventuelles différences entre catcodes de tokens, ce qui devrait permettre de couvrir tous les besoins en matière de programmation.

Table des matières

1	Présentation	2
1.1	Description	2
1.2	Motivation	2
2	Les macros	2
2.1	Les tests	3
2.1.1	<code>\IfSubStr</code>	3
2.1.2	<code>\IfSubStrBefore</code>	3
2.1.3	<code>\IfSubStrBehind</code>	3
2.1.4	<code>\IfBeginWith</code>	4
2.1.5	<code>\IfEndWith</code>	4
2.1.6	<code>\IfInteger</code>	4
2.1.7	<code>\IfDecimal</code>	5
2.1.8	<code>\IfStrEq</code>	5
2.1.9	<code>\IfEq</code>	5
2.1.10	<code>\IfStrEqCase</code>	6
2.1.11	<code>\IfEqCase</code>	6
2.2	Les macros renvoyant une chaîne	6
2.2.1	<code>\StrBefore</code>	6
2.2.2	<code>\StrBehind</code>	7
2.2.3	<code>\StrCut</code>	7
2.2.4	<code>\StrBetween</code>	8
2.2.5	<code>\StrSubstitute</code>	8
2.2.6	<code>\StrDel</code>	8
2.2.7	<code>\StrGobbleLeft</code>	9
2.2.8	<code>\StrLeft</code>	9
2.2.9	<code>\StrGobbleRight</code>	9
2.2.10	<code>\StrRight</code>	9
2.2.11	<code>\StrChar</code>	10
2.2.12	<code>\StrMid</code>	10
2.3	Les macros renvoyant des nombres	10
2.3.1	<code>\StrLen</code>	10
2.3.2	<code>\StrCount</code>	10
2.3.3	<code>\StrPosition</code>	11
2.3.4	<code>\StrCompare</code>	11
3	Modes de fonctionnement	12
3.1	Développement des arguments	12
3.1.1	Les macros <code>\fullexpandarg</code> , <code>\expandarg</code> et <code>\noexpandarg</code>	12
3.1.2	Caractères et lexèmes autorisés dans les arguments	12
3.2	Développement des macros, argument optionnel	13
3.3	Traitement des arguments	13
3.3.1	Traitement à l'unité syntaxique près	13
3.3.2	Exploration des groupes	14
3.4	Catcodes et macros étoilées	14
4	Macros avancées pour la programmation	15
4.1	Recherche d'un groupe, les macros <code>\StrFindGroup</code> et <code>\groupID</code>	15
4.2	Coupe d'une chaîne, la macro <code>\StrSplit</code>	16
4.3	Assigner un contenu verb, la macro <code>\verbtocs</code>	17
4.4	Tokenisation d'un texte vers une séquence de contrôle, la macro <code>\tokenize</code>	18
4.5	Développement contrôlé, les macros <code>\StrExpand</code> et <code>\scancs</code>	18
4.6	À l'intérieur d'une définition de macro	19
4.7	La macro <code>\StrRemoveBraces</code>	19
4.8	Exemples d'utilisation en programmation	20
4.8.1	Exemple 1	20
4.8.2	Exemple 2	20
4.8.3	Exemple 3	21
4.8.4	Exemple 4	21
4.8.5	Exemple 5	21

1 Présentation

1.1 Description

Cette extension¹ regroupe des macros et des tests opérant sur des chaînes de tokens, un peu comme en disposent des langages dit « évolués ». On y trouve les opérations habituelles sur les chaînes, comme par exemple : test si une chaîne en contient une autre, commence ou finit par une autre, test si une chaîne est un nombre entier ou décimal, extractions de sous-chaînes, calculs de position d'une sous-chaîne, calculs du nombre d'occurrences, etc.

On appelle « chaîne de tokens » une suite de tokens quelconques, sachant qu'aucune supposition n'a été faite quant à leur nature, mis à part que dans les chaînes de tokens, les accolades doivent être équilibrées et que les tokens de catcode 6 et 14 (habituellement % et #) n'y sont pas admis. Tous les autres tokens, y compris `\par`, sont *à priori* permis dans n'importe quel ordre, quel que soit le code qu'ils représentent.

Les arguments contenant des chaînes de tokens sont lus par `xstring` *unité syntaxique par unité syntaxique*², ce qui revient à les lire caractère par caractère lorsque ceux-ci contiennent des tokens « normaux », c'est-à-dire dont les catcodes sont 10, 11 et 12. On peut également utiliser `xstring` à des fins de programmation en utilisant des arguments contenant des séquences de contrôle et des tokens dont les catcodes sont moins inoffensifs. Voir le chapitre sur le mode de lecture et de développement des arguments (page 13), la commande `\verbtocs` (page 17), la commande `\scancs` (page 18).

Certes d'autres packages manipulant les chaînes de caractères existent (par exemple `substr` et `stringstrings`), mais outre des différences notables quant aux fonctionnalités, ils ne prennent pas en charge les occurrences des sous-chaînes et me semblent soit trop limités, soit trop difficiles à utiliser pour la programmation.

Comme les macros manipulent des chaînes de tokens, il peut arriver aux utilisateurs avancés de rencontrer des problèmes de « catcodes³ » conduisant à des comportements inattendus. Ces effets indésirables peuvent être contrôlés. Consulter en particulier le chapitre sur les catcodes des arguments page 14.

1.2 Motivation

J'ai été conduit à écrire ce type de macros car je n'ai jamais vraiment trouvé de d'outils sous \LaTeX adaptés à mes besoins concernant le traitement de chaînes. Alors, au fil des mois, et avec l'aide de contributeurs⁴ de `fr.comp.text.tex`, j'ai écrit quelques macros qui me servaient ponctuellement ou régulièrement. Leur nombre s'étant accru, et celles-ci devenant un peu trop dispersées dans les répertoires de mon ordinateur, je les ai regroupées dans ce package.

Ainsi, le fait de donner corps à un ensemble cohérent de macros force à davantage de rigueur et induit naturellement de nécessaires améliorations, ce qui a pris la majeure partie du temps que j'ai consacré à ce package. Pour harmoniser le tout, mais à contre-cœur, j'ai fini par choisir des noms de macros à consonances anglo-saxonnes.

Ensuite, et cela a été ma principale motivation puisque j'ai découvert \LaTeX récemment⁵, l'écriture de `xstring` qui est mon premier package m'a surtout permis de beaucoup progresser en programmation pure, et aborder des méthodes propres à la programmation sous \TeX .

2 Les macros

Pour bien comprendre les actions de chaque macro, envisageons tout d'abord le fonctionnement et la présentation des macros dans leur mode de fonctionnement le plus simple. Pas de problème de catcode ici, ni de tokens spéciaux et encore moins de séquence de contrôle dans les arguments : les arguments contiendront des caractères alphanumériques.

Dans ce chapitre, la totalité des macros est présentée selon ce plan :

- la syntaxe complète⁶ ainsi que la valeur d'éventuels arguments optionnels ;
- une brève description du fonctionnement ;
- le fonctionnement sous certaines conditions particulières. Pour chaque conditions envisagée, le fonctionnement décrit est prioritaire sur celui (ceux) se trouvant au dessous de lui ;

1. L'extension ne nécessite pas \LaTeX et peut être compilée sous Plain $\epsilon\text{-}\TeX$.

2. Sauf cas particulier, une unité syntaxique est un caractère lu dans le code à ces exceptions près : une séquence de contrôle est une unité syntaxique, un groupe entre accolades est aussi une unité syntaxique. Voir également page 13.

3. Codes de catégories, en français.

4. Je remercie chaleureusement Manuel alias « mpg » pour son aide précieuse, sa compétence et sa disponibilité.

5. En novembre 2007, je suis donc un « noob » pour longtemps encore !

6. L'étoile optionnelle après le nom de la macro, et l'argument optionnel entre crochet venant en dernier seront expliqués plus tard. Voir page 14 pour les macros étoilées et page 13 pour l'argument optionnel en dernière position.

- enfin, quelques exemples sont donnés. J’ai essayé de les trouver les plus facilement compréhensibles et les plus représentatifs des situations rencontrées dans une utilisation normale⁷. Si un doute est possible quant à la présence d’espaces dans le résultat, celui-ci sera délimité par des « | », étant entendu qu’une chaîne vide est représentée par « || ».

Important : dans les macros qui suivent, sauf cas spécifié, un $\langle nombre \rangle$ est un nombre entier, un compteur, ou le résultat d’une opération arithmétique effectuée à l’aide de la primitive `\numexpr`.

Dans le reste du texte, les macros de `xstring` sont affichées en rouge.

2.1 Les tests

2.1.1 `\IfSubStr`

`\IfSubStr` $\langle [*] \rangle \langle nombre \rangle \{ \langle chaîne \rangle \} \{ \langle chaîneA \rangle \} \{ \langle vrai \rangle \} \{ \langle faux \rangle \}$

L’argument optionnel $\langle nombre \rangle$ vaut 1 par défaut.

Teste si $\langle chaîne \rangle$ contient au moins $\langle nombre \rangle$ fois $\langle chaîneA \rangle$ et exécute $\langle vrai \rangle$ dans l’affirmative, et $\langle faux \rangle$ dans le cas contraire.

- ▷ Si $\langle nombre \rangle \leq 0$, exécute $\langle faux \rangle$;
- ▷ Si $\langle chaîne \rangle$ ou $\langle chaîneA \rangle$ est vide, exécute $\langle faux \rangle$.

1	<code>\IfSubStr{xstring}{tri}{vrai}{faux}</code>	vrai
2	<code>\IfSubStr{xstring}{a}{vrai}{faux}</code>	faux
3	<code>\IfSubStr{a bc def }{c d}{vrai}{faux}</code>	vrai
4	<code>\IfSubStr{a bc def }{cd}{vrai}{faux}</code>	faux
5	<code>\IfSubStr[2]{1a2a3a}{a}{vrai}{faux}</code>	vrai
6	<code>\IfSubStr[3]{1a2a3a}{a}{vrai}{faux}</code>	vrai
7	<code>\IfSubStr[4]{1a2a3a}{a}{vrai}{faux}</code>	faux

2.1.2 `\IfSubStrBefore`

`\IfSubStrBefore` $\langle [*] \rangle \langle nombre1 \rangle, \langle nombre2 \rangle \{ \langle chaîne \rangle \} \{ \langle chaîneA \rangle \} \{ \langle chaîneB \rangle \} \{ \langle vrai \rangle \} \{ \langle faux \rangle \}$

Les arguments optionnels $\langle nombre1 \rangle$ et $\langle nombre2 \rangle$ valent 1 par défaut.

Dans $\langle chaîne \rangle$, la macro teste si l’occurrence n° $\langle nombre1 \rangle$ de $\langle chaîneA \rangle$ se trouve à gauche de l’occurrence n° $\langle nombre2 \rangle$ de $\langle chaîneB \rangle$. Exécute $\langle vrai \rangle$ dans l’affirmative, et $\langle faux \rangle$ dans le cas contraire.

- ▷ Si l’une des occurrences n’est pas trouvée, exécute $\langle faux \rangle$;
- ▷ Si l’un des arguments $\langle chaîne \rangle$, $\langle chaîneA \rangle$ ou $\langle chaîneB \rangle$ est vide, exécute $\langle faux \rangle$;
- ▷ Si l’un au moins des deux arguments optionnels est négatif ou nul, exécute $\langle faux \rangle$.

1	<code>\IfSubStrBefore{xstring}{st}{in}{vrai}{faux}</code>	vrai
2	<code>\IfSubStrBefore{xstring}{ri}{s}{vrai}{faux}</code>	faux
3	<code>\IfSubStrBefore{LaTeX}{LaT}{TeX}{vrai}{faux}</code>	faux
4	<code>\IfSubStrBefore{a bc def }{ b}{ef}{vrai}{faux}</code>	vrai
5	<code>\IfSubStrBefore{a bc def }{ab}{ef}{vrai}{faux}</code>	faux
6	<code>\IfSubStrBefore[2,1]{b1b2b3}{b}{2}{vrai}{faux}</code>	vrai
7	<code>\IfSubStrBefore[3,1]{b1b2b3}{b}{2}{vrai}{faux}</code>	faux
8	<code>\IfSubStrBefore[2,2]{baobab}{a}{b}{vrai}{faux}</code>	faux
9	<code>\IfSubStrBefore[2,3]{baobab}{a}{b}{vrai}{faux}</code>	vrai

2.1.3 `\IfSubStrBehind`

`\IfSubStrBehind` $\langle [*] \rangle \langle nombre1 \rangle, \langle nombre2 \rangle \{ \langle chaîne \rangle \} \{ \langle chaîneA \rangle \} \{ \langle chaîneB \rangle \} \{ \langle vrai \rangle \} \{ \langle faux \rangle \}$

Les arguments optionnels $\langle nombre1 \rangle$ et $\langle nombre2 \rangle$ valent 1 par défaut.

Dans $\langle chaîne \rangle$, la macro teste si l’occurrence n° $\langle nombre1 \rangle$ de $\langle chaîneA \rangle$ se trouve après l’occurrence n° $\langle nombre2 \rangle$ de $\langle chaîneB \rangle$. Exécute $\langle vrai \rangle$ dans l’affirmative, et $\langle faux \rangle$ dans le cas contraire.

- ▷ Si l’une des occurrences n’est pas trouvée, exécute $\langle faux \rangle$;
- ▷ Si l’un des arguments $\langle chaîne \rangle$, $\langle chaîneA \rangle$ ou $\langle chaîneB \rangle$ est vide, exécute $\langle faux \rangle$;

⁷. Pour une collection plus importante d’exemples, on peut aussi consulter le fichier de test.

▷ Si l'un au moins des deux arguments optionnels est négatif ou nul, exécute $\langle faux \rangle$.

1	<code>\IfSubStrBehind{xstring}{ri}{xs}{vrai}{faux}</code>	faux
2	<code>\IfSubStrBehind{xstring}{s}{i}{vrai}{faux}</code>	faux
3	<code>\IfSubStrBehind{LaTeX}{TeX}{LaT}{vrai}{faux}</code>	faux
4	<code>\IfSubStrBehind{a bc def }{ d}{a}{vrai}{faux}</code>	faux
5	<code>\IfSubStrBehind{a bc def }{cd}{a b}{vrai}{faux}</code>	faux
6	<code>\IfSubStrBehind[2,1]{b1b2b3}{b}{2}{vrai}{faux}</code>	faux
7	<code>\IfSubStrBehind[3,1]{b1b2b3}{b}{2}{vrai}{faux}</code>	vrai
8	<code>\IfSubStrBehind[2,2]{baobab}{b}{a}{vrai}{faux}</code>	faux
9	<code>\IfSubStrBehind[2,3]{baobab}{b}{a}{vrai}{faux}</code>	faux

2.1.4 \IfBeginWith

`\IfBeginWith` $\langle [*] \rangle \{ \langle chaine \rangle \} \{ \langle chaineA \rangle \} \{ \langle vrai \rangle \} \{ \langle faux \rangle \}$

Teste si $\langle chaine \rangle$ commence par $\langle chaineA \rangle$, et exécute $\langle vrai \rangle$ dans l'affirmative, et $\langle faux \rangle$ dans le cas contraire.

▷ Si $\langle chaine \rangle$ ou $\langle chaineA \rangle$ est vide, exécute $\langle faux \rangle$.

1	<code>\IfBeginWith{xstring}{xst}{vrai}{faux}</code>	vrai
2	<code>\IfBeginWith{LaTeX}{a}{vrai}{faux}</code>	faux
3	<code>\IfBeginWith{a bc def }{a b}{vrai}{faux}</code>	vrai
4	<code>\IfBeginWith{a bc def }{ab}{vrai}{faux}</code>	faux

2.1.5 \IfEndWith

`\IfEndWith` $\langle [*] \rangle \{ \langle chaine \rangle \} \{ \langle chaineA \rangle \} \{ \langle vrai \rangle \} \{ \langle faux \rangle \}$

Teste si $\langle chaine \rangle$ se termine par $\langle chaineA \rangle$, et exécute $\langle vrai \rangle$ dans l'affirmative, et $\langle faux \rangle$ dans le cas contraire.

▷ Si $\langle chaine \rangle$ ou $\langle chaineA \rangle$ est vide, exécute $\langle faux \rangle$.

1	<code>\IfEndWith{xstring}{ring}{vrai}{faux}</code>	vrai
2	<code>\IfEndWith{LaTeX}{a}{vrai}{faux}</code>	faux
3	<code>\IfEndWith{a bc def }{ef }{vrai}{faux}</code>	vrai
4	<code>\IfEndWith{a bc def }{ef}{vrai}{faux}</code>	faux

2.1.6 \IfInteger

`\IfInteger` $\{ \langle nombre \rangle \} \{ \langle vrai \rangle \} \{ \langle faux \rangle \}$

Teste si $\langle nombre \rangle$ est un nombre entier relatif (c'est-à-dire dont la partie décimale est absente ou constituée d'une suite de 0), et exécute $\langle vrai \rangle$ dans l'affirmative, et $\langle faux \rangle$ dans le cas contraire.

Si le test est faux pour cause de caractères non autorisés, les séquences de contrôle `\@xs@afterinteger` et `\afterinteger` contiennent chacune la partie illégale de $\langle nombre \rangle$.

1	<code>\IfInteger{13}{vrai}{faux}</code>	vrai
2	<code>\IfInteger{-219}{vrai}{faux}</code>	vrai
3	<code>\IfInteger{+9}{vrai}{faux}</code>	vrai
4	<code>\IfInteger{3.14}{vrai}{faux}</code>	faux
5	<code>\IfInteger{8.0}{vrai}{faux}</code>	vrai
6	<code>\IfInteger{0}{vrai}{faux}</code>	vrai
7	<code>\IfInteger{49a}{vrai}{faux}</code>	faux
8	<code>\IfInteger{+}{vrai}{faux}</code>	faux
9	<code>\IfInteger{-}{vrai}{faux}</code>	faux
10	<code>\IfInteger{0000}{vrai}{faux}</code>	vrai

2.1.7 \IfDecimal

`\IfDecimal{⟨nombre⟩}{⟨vrai⟩}{⟨faux⟩}`

Teste si `⟨nombre⟩` est un nombre décimal, et exécute `⟨vrai⟩` dans l'affirmative, et `⟨faux⟩` dans le cas contraire.

Les compteurs `\integerpart` et `\decimalpart` contiennent les parties entières et décimales de `⟨nombre⟩`.

Si le test est faux pour cause de caractères non autorisés, les séquences de contrôle `\@xs@afterdecimal` et `\afterdecimal` contiennent chacune la partie illégale de `⟨nombre⟩`, alors que si le test est faux parce que la partie décimale après le séparateur décimal est vide, elles contiennent « X ».

- ▷ Le séparateur décimal peut être un point ou une virgule;
- ▷ Si le `⟨nombre⟩` commence par un séparateur décimal, le test est vrai (lignes 4 et 5) et la partie entière est considérée comme étant 0;
- ▷ Si le `⟨nombre⟩` se termine par un séparateur décimal, le test est faux (lignes 9 et 10).

1	<code>\IfDecimal{3.14}{vrai}{faux}</code>	vrai
2	<code>\IfDecimal{3,14}{vrai}{faux}</code>	vrai
3	<code>\IfDecimal{-0.5}{vrai}{faux}</code>	vrai
4	<code>\IfDecimal{.7}{vrai}{faux}</code>	vrai
5	<code>\IfDecimal{,9}{vrai}{faux}</code>	vrai
6	<code>\IfDecimal{1..2}{vrai}{faux}</code>	faux
7	<code>\IfDecimal{+6}{vrai}{faux}</code>	vrai
8	<code>\IfDecimal{-15}{vrai}{faux}</code>	vrai
9	<code>\IfDecimal{1.}{vrai}{faux}</code>	faux
10	<code>\IfDecimal{2,}{vrai}{faux}</code>	faux
11	<code>\IfDecimal{.}{vrai}{faux}</code>	faux
12	<code>\IfDecimal{,}{vrai}{faux}</code>	faux
13	<code>\IfDecimal{+}{vrai}{faux}</code>	faux
14	<code>\IfDecimal{-}{vrai}{faux}</code>	faux

2.1.8 \IfStrEq

`\IfStrEq[*]{⟨chaîneA⟩}{⟨chaîneB⟩}{⟨vrai⟩}{⟨faux⟩}`

Teste si les chaînes `⟨chaîneA⟩` et `⟨chaîneB⟩` sont égales, c'est-à-dire si elles contiennent successivement les mêmes unités syntaxiques dans le même ordre. Exécute `⟨vrai⟩` dans l'affirmative, et `⟨faux⟩` dans le cas contraire.

1	<code>\IfStrEq{a1b2c3}{a1b2c3}{vrai}{faux}</code>	vrai
2	<code>\IfStrEq{abcdef}{abcd}{vrai}{faux}</code>	faux
3	<code>\IfStrEq{abc}{abcdef}{vrai}{faux}</code>	faux
4	<code>\IfStrEq{3,14}{3,14}{vrai}{faux}</code>	vrai
5	<code>\IfStrEq{12.34}{12.340}{vrai}{faux}</code>	faux
6	<code>\IfStrEq{abc}{}{vrai}{faux}</code>	faux
7	<code>\IfStrEq{}{abc}{vrai}{faux}</code>	faux
8	<code>\IfStrEq{}{}{vrai}{faux}</code>	vrai

2.1.9 \IfEq

`\IfEq[*]{⟨chaîneA⟩}{⟨chaîneB⟩}{⟨vrai⟩}{⟨faux⟩}`

Teste si les chaînes `⟨chaîneA⟩` et `⟨chaîneB⟩` sont égales, *sauf* si `⟨chaîneA⟩` et `⟨chaîneB⟩` contiennent des nombres, auquel cas la macro teste si les nombres sont égaux. Exécute `⟨vrai⟩` dans l'affirmative, et `⟨faux⟩` dans le cas contraire.

- ▷ La définition de *nombre* est celle évoquée dans la macro `\IfDecimal` (voir page 5), et donc :
- ▷ Les signes « + » sont facultatifs;
- ▷ Le séparateur décimal peut être indifféremment la virgule ou le point.

1	<code>\IfEq{a1b2c3}{a1b2c3}{vrai}{faux}</code>	vrai
2	<code>\IfEq{abcdef}{ab}{vrai}{faux}</code>	faux
3	<code>\IfEq{ab}{abcdef}{vrai}{faux}</code>	faux
4	<code>\IfEq{12.34}{12,34}{vrai}{faux}</code>	vrai
5	<code>\IfEq{+12.34}{12.340}{vrai}{faux}</code>	vrai
6	<code>\IfEq{10}{+10}{vrai}{faux}</code>	vrai
7	<code>\IfEq{-10}{10}{vrai}{faux}</code>	faux
8	<code>\IfEq{+0,5}{,5}{vrai}{faux}</code>	vrai
9	<code>\IfEq{1.001}{1.01}{vrai}{faux}</code>	faux
10	<code>\IfEq{3*4+2}{14}{vrai}{faux}</code>	faux
11	<code>\IfEq{\number\numexpr3*4+2}{14}{vrai}{faux}</code>	vrai
12	<code>\IfEq{0}{-0.0}{vrai}{faux}</code>	vrai
13	<code>\IfEq{}{}{vrai}{faux}</code>	vrai

2.1.10 \IfStrEqCase

```
\IfStrEqCase⟨[*]⟩{⟨chaîne⟩}{%
  {⟨chaîne1⟩}{⟨code1⟩}%
  {⟨chaîne2⟩}{⟨code2⟩}%
  etc...
  {⟨chaîneN⟩}{⟨codeN⟩}}[⟨code alternatif⟩]
```

Teste successivement si *⟨chaîne⟩* est égale à *⟨chaîne1⟩*, *⟨chaîne2⟩*, etc. La comparaison se fait au sens de `\IfStrEq` (voir paragraphes précédents). Si un test est positif, le code correspondant est exécuté et la macro se termine. Si tous les tests sont négatifs, le code optionnel *⟨code alternatif⟩* est exécuté s'il est présent.

1	<code>\IfStrEqCase{b}{a}{AA}{b}{BB}{c}{CC}</code>	BB
2	<code> \IfStrEqCase{abc}{a}{AA}{b}{BB}{c}{CC} </code>	
3	<code>\IfStrEqCase{c}{a}{AA}{b}{BB}{c}{CC}[autre]</code>	CC
4	<code>\IfStrEqCase{d}{a}{AA}{b}{BB}{c}{CC}[autre]</code>	autre
5	<code>\IfStrEqCase{+3}{1}{un}{2}{deux}{3}{trois}[autre]</code>	autre
6	<code>\IfStrEqCase{0.5}{0}{zero}{.5}{demi}{1}{un}[autre]</code>	autre

2.1.11 \IfEqCase

```
\IfEqCase⟨[*]⟩{⟨chaîne⟩}{%
  {⟨chaîne1⟩}{⟨code1⟩}%
  {⟨chaîne2⟩}{⟨code2⟩}%
  etc...
  {⟨chaîneN⟩}{⟨codeN⟩}}[⟨code alternatif⟩]
```

Teste successivement si *⟨chaîne⟩* est égale à *⟨chaîne1⟩*, *⟨chaîne2⟩*, etc. La comparaison se fait au sens de `\IfEq` (voir paragraphes précédents). Si un test est positif, le code correspondant est exécuté et la macro se termine. Si tous les tests sont négatifs, le code optionnel *⟨code alternatif⟩* est exécuté s'il est présent.

1	<code>\IfEqCase{b}{a}{AA}{b}{BB}{c}{CC}</code>	BB
2	<code> \IfEqCase{abc}{a}{AA}{b}{BB}{c}{CC} </code>	
3	<code>\IfEqCase{c}{a}{AA}{b}{BB}{c}{CC}[autre]</code>	CC
4	<code>\IfEqCase{d}{a}{AA}{b}{BB}{c}{CC}[autre]</code>	autre
5	<code>\IfEqCase{+3}{1}{un}{2}{deux}{3}{trois}[autre]</code>	trois
6	<code>\IfEqCase{0.5}{0}{zero}{.5}{demi}{1}{un}[autre]</code>	demi

2.2 Les macros renvoyant une chaîne

2.2.1 \StrBefore

```
\StrBefore⟨[*]⟩[⟨nombre⟩]{⟨chaîne⟩}{⟨chaîneA⟩}[⟨nom⟩]
```

L'argument optionnel *⟨nombre⟩* vaut 1 par défaut.

Dans *⟨chaîne⟩*, renvoie ce qui se trouve avant l'occurrence n° *⟨nombre⟩* de *⟨chaîneA⟩*.

- ▷ Si $\langle chaîne \rangle$ ou $\langle chaîneA \rangle$ est vide, une chaîne vide est renvoyée ;
- ▷ Si $\langle nombre \rangle < 1$ alors, la macro se comporte comme si $\langle nombre \rangle = 1$;
- ▷ Si l'occurrence n'est pas trouvée, une chaîne vide est renvoyée.

<pre> 1 \StrBefore{xstring}{tri} 2 \StrBefore{LaTeX}{e} 3 \StrBefore{LaTeX}{p} 4 \StrBefore{LaTeX}{L} 5 \StrBefore{a bc def }{def} 6 \StrBefore{a bc def }{cd} 7 \StrBefore[1]{1b2b3}{b} 8 \StrBefore[2]{1b2b3}{b} </pre>	<pre> xs LaT a bc 1 1b2 </pre>
---	--

2.2.2 \StrBehind

$\backslash\text{StrBehind}\langle [*] \rangle [\langle nombre \rangle] \{ \langle chaîne \rangle \} \{ \langle chaîneA \rangle \} [\langle nom \rangle]$

L'argument optionnel $\langle nombre \rangle$ vaut 1 par défaut.

Dans $\langle chaîne \rangle$, renvoie ce qui se trouve après l'occurrence n° $\langle nombre \rangle$ de $\langle chaîneA \rangle$.

- ▷ Si $\langle chaîne \rangle$ ou $\langle chaîneA \rangle$ est vide, une chaîne vide est renvoyée ;
- ▷ Si $\langle nombre \rangle < 1$ alors, la macro se comporte comme si $\langle nombre \rangle = 1$;
- ▷ Si l'occurrence n'est pas trouvée, une chaîne vide est renvoyée.

<pre> 1 \StrBehind{xstring}{tri} 2 \StrBehind{LaTeX}{e} 3 \StrBehind{LaTeX}{p} 4 \StrBehind{LaTeX}{X} 5 \StrBehind{a bc def }{bc} 6 \StrBehind{a bc def }{cd} 7 \StrBehind[1]{1b2b3}{b} 8 \StrBehind[2]{1b2b3}{b} 9 \StrBehind[3]{1b2b3}{b} </pre>	<pre> ng X def 2b3 3 </pre>
--	---

2.2.3 \StrCut

Voici la syntaxe de cette macro :

$\backslash\text{StrCut}\langle [*] \rangle [\langle nombre \rangle] \{ \langle chaîne \rangle \} \{ \langle chaîneA \rangle \} \{ \langle macroA \rangle \} \{ \langle macroB \rangle \}$

L'argument optionnel $\langle nombre \rangle$ vaut 1 par défaut.

La $\langle chaîne \rangle$ est coupée en deux chaînes à l'occurrence n° $[\langle nombre \rangle]$ de la $\{ \langle chaîneA \rangle \}$. Ce qui se trouve à gauche de cette occurrence est assigné à la séquence de contrôle $\langle macroA \rangle$ et ce qui se trouve à droite à $\langle macroB \rangle$.

Cette macro renvoie *deux chaînes* et donc **n'affiche rien**. Par conséquent, elle ne dispose pas de l'argument optionnel en dernière position.

- ▷ Si $\langle chaîne \rangle$ ou $\langle chaîneA \rangle$ est vide, $\langle macroA \rangle$ et $\langle macroB \rangle$ seront vides ;
- ▷ Si $\langle nombre \rangle < 1$ alors, la macro se comporte comme si $\langle nombre \rangle = 1$;
- ▷ Si l'occurrence n'est pas trouvée, $\langle macroA \rangle$ reçoit la totalité de $\langle chaîne \rangle$ tandis que $\langle macroB \rangle$ est vide.

<pre> 1 \StrCut{abracadabra}{a}\csA\csB \csA \csB \par 2 \StrCut[2]{abracadabra}{a}\csA\csB \csA \csB \par 3 \StrCut[3]{abracadabra}{a}\csA\csB \csA \csB \par 4 \StrCut[4]{abracadabra}{a}\csA\csB \csA \csB \par 5 \StrCut[5]{abracadabra}{a}\csA\csB \csA \csB \par 6 \StrCut[6]{abracadabra}{a}\csA\csB \csA \csB \par 7 \StrCut[-4]{abracadabra}{a}\csA\csB \csA \csB \par 8 \StrCut{abracadabra}{brac}\csA\csB \csA \csB \par 9 \StrCut{abracadabra}{foo}\csA\csB \csA \csB \par 10 \StrCut{abracadabra}{}\csA\csB \csA \csB </pre>	<pre> bracadabra abr cadabra abrac dabra abracad bra abracadabr abracadabra bracadabra a adabra abracadabra </pre>
--	---

2.2.4 \StrBetween

`\StrBetween`[[*]] [⟨nombre1⟩, ⟨nombre2⟩] {⟨chaîne⟩} {⟨chaîneA⟩} {⟨chaîneB⟩} [⟨nom⟩]

Les arguments optionnels ⟨nombre1⟩ et ⟨nombre2⟩ valent 1 par défaut.

Dans ⟨chaîne⟩, renvoie ce qui se trouve entre⁸ les occurrences n° ⟨nombre1⟩ de ⟨chaîneA⟩ et n° ⟨nombre2⟩ de ⟨chaîneB⟩.

- ▷ Si les occurrences ne sont pas dans l'ordre (⟨chaîneA⟩ puis ⟨chaîneB⟩) dans ⟨chaîne⟩, une chaîne vide est renvoyée;
- ▷ Si l'une des 2 occurrences n'existe pas dans ⟨chaîne⟩, une chaîne vide est renvoyée;
- ▷ Si l'un des arguments optionnels ⟨nombre1⟩ ou ⟨nombre2⟩ est négatif ou nul, une chaîne vide est renvoyée.

1	<code>\StrBetween{xstring}{xs}{ng}</code>	tri
2	<code> \StrBetween{xstring}{i}{n} </code>	
3	<code> \StrBetween{xstring}{a}{tring} </code>	
4	<code> \StrBetween{a bc def }{a}{d} </code>	bc
5	<code> \StrBetween{a bc def }{a }{f} </code>	bc de
6	<code>\StrBetween{a1b1a2b2a3b3}{a}{b}</code>	1
7	<code>\StrBetween[2,3]{a1b1a2b2a3b3}{a}{b}</code>	2b2a3
8	<code>\StrBetween[1,3]{a1b1a2b2a3b3}{a}{b}</code>	1b1a2b2a3
9	<code> \StrBetween[3,1]{a1b1a2b2a3b3}{a}{b} </code>	
10	<code>\StrBetween[3,2]{abracadabra}{a}{bra}</code>	da

2.2.5 \StrSubstitute

`\StrSubstitute`[⟨nombre⟩] {⟨chaîne⟩} {⟨chaîneA⟩} {⟨chaîneB⟩} [⟨nom⟩]

L'argument optionnel ⟨nombre⟩ vaut 0 par défaut.

Dans ⟨chaîne⟩, la macro remplace les ⟨nombre⟩ premières occurrences de ⟨chaîneA⟩ par ⟨chaîneB⟩, sauf si ⟨nombre⟩ = 0 auquel cas, toutes les occurrences sont remplacées.

- ▷ Si ⟨chaîne⟩ est vide, une chaîne vide est renvoyée;
- ▷ Si ⟨chaîneA⟩ est vide ou n'existe pas dans ⟨chaîne⟩, la macro est sans effet;
- ▷ Si ⟨nombre⟩ est supérieur au nombre d'occurrences de ⟨chaîneA⟩, alors toutes les occurrences sont remplacées;
- ▷ Si ⟨nombre⟩ < 0 alors la macro se comporte comme si ⟨nombre⟩ = 0;
- ▷ Si ⟨chaîneB⟩ est vide, alors les occurrences de ⟨chaîneA⟩, si elles existent, sont supprimées.

1	<code>\StrSubstitute{xstring}{i}{a}</code>	xstrang
2	<code>\StrSubstitute{abracadabra}{a}{o}</code>	obrocodobro
3	<code>\StrSubstitute{abracadabra}{br}{TeX}</code>	aTeXacadaTeXa
4	<code>\StrSubstitute{LaTeX}{m}{n}</code>	LaTeX
5	<code>\StrSubstitute{a bc def }{ }{M}</code>	aMbcMdefM
6	<code>\StrSubstitute{a bc def }{ab}{AB}</code>	a bc def
7	<code>\StrSubstitute[1]{a1a2a3}{a}{B}</code>	B1a2a3
8	<code>\StrSubstitute[2]{a1a2a3}{a}{B}</code>	B1B2a3
9	<code>\StrSubstitute[3]{a1a2a3}{a}{B}</code>	B1B2B3
10	<code>\StrSubstitute[4]{a1a2a3}{a}{B}</code>	B1B2B3

2.2.6 \StrDel

`\StrDel`[⟨nombre⟩] {⟨chaîne⟩} {⟨chaîneA⟩} [⟨nom⟩]

L'argument optionnel ⟨nombre⟩ vaut 0 par défaut.

Supprime les ⟨nombre⟩ premières occurrences de ⟨chaîneA⟩ dans ⟨chaîne⟩, sauf si ⟨nombre⟩ = 0 auquel cas, toutes les occurrences sont supprimées.

- ▷ Si ⟨chaîne⟩ est vide, une chaîne vide est renvoyée;
- ▷ Si ⟨chaîneA⟩ est vide ou n'existe pas dans ⟨chaîne⟩, la macro est sans effet;
- ▷ Si ⟨nombre⟩ est supérieur au nombre d'occurrences de ⟨chaîneA⟩, alors toutes les occurrences sont supprimées;
- ▷ Si ⟨nombre⟩ < 0 alors la macro se comporte comme si ⟨nombre⟩ = 0;

8. Au sens strict, c'est-à-dire sans les chaînes frontière

<pre> 1 \StrDel{abracadabra}{a} 2 \StrDel[1]{abracadabra}{a} 3 \StrDel[4]{abracadabra}{a} 4 \StrDel[9]{abracadabra}{a} 5 \StrDel{a bc def }{ } 6 \StrDel{a bc def }{def} </pre>	<pre> brcdr bracadabra brcdbra brcdr abcdef a bc </pre>
---	--

2.2.7 \StrGobbleLeft

`\StrGobbleLeft{<chaîne>}{<nombre>}[<nom>]`

Dans *<chaîne>*, enlève les *<nombre>* premières unités syntaxiques de gauche.

- ▷ Si *<chaîne>* est vide, renvoie une chaîne vide ;
- ▷ Si $\langle nombre \rangle \leq 0$, aucune unité syntaxique n'est supprimée ;
- ▷ Si $\langle nombre \rangle \geq \langle longueurChaîne \rangle$, toutes les unités syntaxiques sont supprimées.

<pre> 1 \StrGobbleLeft{xstring}{2} 2 \StrGobbleLeft{xstring}{9} 3 \StrGobbleLeft{LaTeX}{4} 4 \StrGobbleLeft{LaTeX}{-2} 5 \StrGobbleLeft{a bc def }{4} </pre>	<pre> tring X LaTeX def </pre>
--	-------------------------------------

2.2.8 \StrLeft

`\StrLeft{<chaîne>}{<nombre>}[<nom>]`

Dans *<chaîne>*, renvoie la sous-chaîne de gauche de longueur *<nombre>*.

- ▷ Si *<chaîne>* est vide, renvoie une chaîne vide ;
- ▷ Si $\langle nombre \rangle \leq 0$, aucune unité syntaxique n'est retournée ;
- ▷ Si $\langle nombre \rangle \geq \langle longueurChaîne \rangle$, toutes les unités syntaxiques sont retournées.

<pre> 1 \StrLeft{xstring}{2} 2 \StrLeft{xstring}{9} 3 \StrLeft{LaTeX}{4} 4 \StrLeft{LaTeX}{-2} 5 \StrLeft{a bc def }{5} </pre>	<pre> xs xstring LaTe a bc </pre>
--	---

2.2.9 \StrGobbleRight

`\StrGobbleRight{<chaîne>}{<nombre>}[<nom>]`

Agit comme `\StrGobbleLeft`, mais enlève les unités syntaxiques à droite de *<chaîne>*.

<pre> 1 \StrGobbleRight{xstring}{2} 2 \StrGobbleRight{xstring}{9} 3 \StrGobbleRight{LaTeX}{4} 4 \StrGobbleRight{LaTeX}{-2} 5 \StrGobbleRight{a bc def }{4} </pre>	<pre> xstri L LaTeX a bc </pre>
---	---

2.2.10 \StrRight

`\StrRight{<chaîne>}{<nombre>}[<nom>]`

Agit comme `\StrLeft`, mais renvoie les unités syntaxiques à la droite de *<chaîne>*.

<pre> 1 \StrRight{xstring}{2} 2 \StrRight{xstring}{9} 3 \StrRight{LaTeX}{4} 4 \StrRight{LaTeX}{-2} 5 \StrRight{a bc def }{5} </pre>	<pre> ng xstring aTeX def </pre>
---	-------------------------------------

2.2.11 \StrChar

`\StrChar{⟨chaîne⟩}{⟨nombre⟩}[⟨nom⟩]`

Renvoie l'unité syntaxique à la position `⟨nombre⟩` dans la chaîne `⟨chaîne⟩`.

- ▷ Si `⟨chaîne⟩` est vide, aucune unité syntaxique n'est renvoyée ;
- ▷ Si `⟨nombre⟩ ≤ 0` ou si `⟨nombre⟩ > ⟨longueurChaîne⟩`, aucune unité syntaxique n'est renvoyée.

1	<code>\StrChar{xstring}{4}</code>	r
2	<code> \StrChar{xstring}{9} </code>	
3	<code> \StrChar{xstring}{-5} </code>	
4	<code>\StrChar{a bc def }{6}</code>	d

2.2.12 \StrMid

`\StrMid{⟨chaîne⟩}{⟨nombre1⟩}{⟨nombre2⟩}[⟨nom⟩]`

Dans `⟨chaîne⟩`, renvoie la sous chaîne se trouvant entre⁹ les positions `⟨nombre1⟩` et `⟨nombre2⟩`.

- ▷ Si `⟨chaîne⟩` est vide, une chaîne vide est renvoyée ;
- ▷ Si `⟨nombre1⟩ > ⟨nombre2⟩`, alors rien n'est renvoyé ;
- ▷ Si `⟨nombre1⟩ < 1` et `⟨nombre2⟩ < 1` alors rien n'est renvoyé ;
- ▷ Si `⟨nombre1⟩ > ⟨longueurChaîne⟩` et `⟨nombre2⟩ > ⟨longueurChaîne⟩`, alors rien n'est renvoyé ;
- ▷ Si `⟨nombre1⟩ < 1`, alors la macro se comporte comme si `⟨nombre1⟩ = 1` ;
- ▷ Si `⟨nombre2⟩ > ⟨longueurChaîne⟩`, alors la macro se comporte comme si `⟨nombre2⟩ = ⟨longueurChaîne⟩`.

1	<code>\StrMid{xstring}{2}{5}</code>	stri
2	<code>\StrMid{xstring}{-4}{2}</code>	xs
3	<code> \StrMid{xstring}{5}{1} </code>	
4	<code>\StrMid{xstring}{6}{15}</code>	ng
5	<code>\StrMid{xstring}{3}{3}</code>	t
6	<code> \StrMid{a bc def }{2}{7} </code>	bc de

2.3 Les macros renvoyant des nombres

2.3.1 \StrLen

`\StrLen{⟨chaîne⟩}[⟨nom⟩]`

Renvoie la longueur de `⟨chaîne⟩`.

1	<code>\StrLen{xstring}</code>	7
2	<code>\StrLen{A}</code>	1
3	<code>\StrLen{a bc def }</code>	9

2.3.2 \StrCount

`\StrCount{⟨chaîne⟩}{⟨chaîneA⟩}[⟨nom⟩]`

Compte combien de fois `⟨chaîneA⟩` est contenue dans `⟨chaîne⟩`.

- ▷ Si l'un au moins des arguments `⟨chaîne⟩` ou `⟨chaîneA⟩` est vide, la macro renvoie 0.

1	<code>\StrCount{abracadabra}{a}</code>	5
2	<code>\StrCount{abracadabra}{bra}</code>	2
3	<code>\StrCount{abracadabra}{tic}</code>	0
4	<code>\StrCount{aaaaaa}{aa}</code>	3

9. Au sens large, c'est-à-dire que les chaînes « frontière » sont renvoyés.

2.3.3 \StrPosition

`\StrPosition[⟨nombre⟩]{⟨chaine⟩}{⟨chaineA⟩}[⟨nom⟩]`

L'argument optionnel `⟨nombre⟩` vaut 1 par défaut.

Dans `⟨chaine⟩`, renvoie la position de l'occurrence n° `⟨nombre⟩` de `⟨chaineA⟩`.

- ▷ Si `⟨nombre⟩` est supérieur au nombre d'occurrences de `⟨chaineA⟩`, alors la macro renvoie 0.
- ▷ Si `⟨chaine⟩` ne contient pas `⟨chaineA⟩`, alors la macro renvoie 0.

```
1 \StrPosition{xstring}{ring} 4
2 \StrPosition[4]{abracadabra}{a} 8
3 \StrPosition[2]{abracadabra}{bra} 9
4 \StrPosition[9]{abracadabra}{a} 0
5 \StrPosition{abracadabra}{z} 0
6 \StrPosition{a bc def }{d} 6
7 \StrPosition[3]{aaaaaa}{aa} 5
```

2.3.4 \StrCompare

`\StrCompare[⟨*⟩]{⟨chaineA⟩}{⟨chaineB⟩}[⟨nom⟩]`

Cette macro peut fonctionner avec 2 tolérances, la tolérance « normale » qui est sélectionnée par défaut et la tolérance « stricte ».

- La tolérance normale, activée par la commande `\comparenormal`.
La macro compare successivement les unités syntaxiques de gauche à droite des chaînes `⟨chaineA⟩` et `⟨chaineB⟩` jusqu'à ce qu'une différence apparaisse ou que la fin de la plus courte chaîne soit atteinte. Si aucune différence n'est trouvée, la macro renvoie 0. Sinon, la position de la 1^{re} différence est renvoyée.
- La tolérance stricte, activée par la commande `\comparestrict`.
La macro compare les 2 chaînes. Si elles sont égales, elle renvoie 0 sinon la position de la 1^{re} différence est renvoyée.

L'ordre des 2 chaînes n'a aucune influence sur le comportement de la macro.

On peut également mémoriser le mode de comparaison en cours avec `\savecomparemode`, le modifier par la suite et revenir à la situation lors de la sauvegarde avec `\restorecomparemode`.

Exemples en tolérance normale :

```
1 \StrCompare{abcd}{abcd} 0
2 \StrCompare{abcd}{abc} 0
3 \StrCompare{abc}{abcd} 0
4 \StrCompare{a b c}{abc} 2
5 \StrCompare{aaa}{baaa} 1
6 \StrCompare{abc}{xyz} 1
7 \StrCompare{123456}{123457} 6
8 \StrCompare{abc}{} 0
```

Exemples en tolérance stricte :

```
1 \StrCompare{abcd}{abcd} 0
2 \StrCompare{abcd}{abc} 4
3 \StrCompare{abc}{abcd} 4
4 \StrCompare{a b c}{abc} 2
5 \StrCompare{aaa}{baaa} 1
6 \StrCompare{abc}{xyz} 1
7 \StrCompare{123456}{123457} 6
8 \StrCompare{abc}{} 1
```

3 Modes de fonctionnement

3.1 Développement des arguments

3.1.1 Les macros `\fullexpandarg`, `\expandarg` et `\noexpandarg`

La macro `\fullexpandarg` est appelée par défaut, ce qui fait que certains arguments (en violet dans la liste ci dessous) transmis aux macros sont développés le plus possible (pour cela, un `\edef` est utilisé). Ce mode de développement maximal permet dans la plupart des cas d'éviter d'utiliser des chaînes d'`\expandafter`. Le code en est souvent allégé.

On peut interdire le développement de ces arguments (et ainsi revenir au comportement normal de \TeX) en invoquant `\noexpandarg` ou `\normalexpandarg` qui sont synonymes.

Il existe enfin un autre mode de développement de ces arguments que l'on appelle avec `\expandarg`. Dans ce cas, le **premier token** de ces arguments est développé *une fois* avant que la macro ne soit appelée. Si l'argument contient plus d'un token, les tokens qui suivent le premier ne sont pas développés (on peut contourner cette volontaire limitation et utiliser la macro `\StrExpand`, voir page 18).

Les commandes `\fullexpandarg`, `\noexpandarg`, `\normalexpandarg` et `\expandarg` peuvent être appelées à tout moment dans le code et fonctionnent comme des bascules. On peut rendre leur portée locale dans un groupe.

On peut également mémoriser le mode de développement en cours avec `\saveexpandmode`, le modifier par la suite et revenir à la situation lors de la sauvegarde avec `\restoreexpandmode`.

Dans la liste ci-dessous, on peut voir en violet quels arguments sont soumis à l'éventuel développement pour chaque macro vue dans le chapitre précédent :

- `\IfSubStr` $\langle[*]\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfSubStrBefore` $\langle[*]\rangle$ $\langle\text{nombre1}\rangle$, $\langle\text{nombre2}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfSubStrBehind` $\langle[*]\rangle$ $\langle\text{nombre1}\rangle$, $\langle\text{nombre2}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfBeginWith` $\langle[*]\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfEndWith` $\langle[*]\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfInteger` $\langle\text{nombre}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfDecimal` $\langle\text{nombre}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfStrEq` $\langle[*]\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfEq` $\langle[*]\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ $\langle\text{vrai}\rangle$ $\langle\text{faux}\rangle$
- `\IfStrEqCase` $\langle[*]\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîne1}\rangle$ $\langle\text{code1}\rangle$
 $\langle\text{chaîne2}\rangle$ $\langle\text{code2}\rangle$
...
 $\langle\text{chaîne } n\rangle$ $\langle\text{code } n\rangle$ $\langle\text{code alternatif}\rangle$
- `\IfEqCase` $\langle[*]\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîne1}\rangle$ $\langle\text{code1}\rangle$
 $\langle\text{chaîne2}\rangle$ $\langle\text{code2}\rangle$
...
 $\langle\text{chaîne } n\rangle$ $\langle\text{code } n\rangle$ $\langle\text{code alternatif}\rangle$
- `\StrBefore` $\langle[*]\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{nom}\rangle$
- `\StrBehind` $\langle[*]\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{nom}\rangle$
- `\StrBetween` $\langle[*]\rangle$ $\langle\text{nombre1}\rangle$, $\langle\text{nombre2}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ $\langle\text{nom}\rangle$
- `\StrSubstitute` $\langle\text{nombre}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ $\langle\text{nom}\rangle$
- `\StrDel` $\langle\text{nombre}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{nom}\rangle$
- `\StrSplit` $\langle\text{chaîne}\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ (voir page 16 pour la macro `StrSplit`)
- `\StrGobbleLeft` $\langle\text{chaîne}\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{nom}\rangle$
- `\StrLeft` $\langle\text{chaîne}\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{nom}\rangle$
- `\StrGobbleRight` $\langle\text{chaîne}\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{nom}\rangle$
- `\StrRight` $\langle\text{chaîne}\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{nom}\rangle$
- `\StrChar` $\langle\text{chaîne}\rangle$ $\langle\text{nombre}\rangle$ $\langle\text{nom}\rangle$
- `\StrMid` $\langle\text{chaîne}\rangle$ $\langle\text{nombre1}\rangle$ $\langle\text{nombre2}\rangle$ $\langle\text{nom}\rangle$
- `\StrLen` $\langle\text{chaîne}\rangle$ $\langle\text{nom}\rangle$
- `\StrCount` $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{nom}\rangle$
- `\StrPosition` $\langle\text{nombre}\rangle$ $\langle\text{chaîne}\rangle$ $\langle\text{chaîneA}\rangle$ $\langle\text{nom}\rangle$
- `\StrCompare` $\langle\text{chaîneA}\rangle$ $\langle\text{chaîneB}\rangle$ $\langle\text{nom}\rangle$

3.1.2 Caractères et lexèmes autorisés dans les arguments

Tout d'abord, quelque soit le mode de développement choisi, **les tokens de catcode 6 et 14 (habituellement # et %)** sont interdits dans tous les arguments ¹⁰.

10. Le token # sera peut-être autorisé dans une future version !

Lorsque le mode `\fullexpandarg` est activé, les arguments sont évalués à l'aide de la primitive `\edef` avant d'être transmis aux macros. Par conséquent, sont autorisés dans les arguments :

- les lettres, majuscules, minuscules, accentuées¹¹ ou non, les chiffres, les espaces¹² ainsi que tout autre token de catcode 10, 11 ou 12 (signes de ponctuation, signes opératoires mathématiques, parenthèses, crochets, etc);
- les tokens de catcode 1 à 4, qui sont habituellement : « { » « } »¹³ « \$ » « & »
- les tokens de catcode 7 et 8, qui sont habituellement : « ^ » « _ »
- toute séquence de contrôle si elle est purement développable¹⁴ et dont le développement maximal donne des caractères autorisés;
- un token de catcode 13 (caractère actif) s'il est purement développable.

Lorsque les arguments ne sont plus développés (utilisation de `\noexpandarg`), on peut aussi inclure dans les arguments n'importe quel token, quelque soit le code qui en résulte comme par exemple toute séquence de contrôle, même non définie ou tout token de catcode 13. On peut également inclure dans les arguments des tokens de test comme `\if` ou `\ifx` ou tout autre token de test, même sans leur `\fi` correspondant; ou bien un `\csname` sans le `\endcsname` correspondant. Dans l'exemple suivant, l'argument contient un `\ifx` sans le `\fi`, et l'on isole ce qui est entre le `\ifx` et le `\else` :

<pre> 1 \noexpandarg 2 \StrBetween{\ifx ab faux \else vrai}{\ifx}{\else} </pre>	ab faux
---	---------

Lorsqu'on utilise `\expandarg`, la précaution concerne le premier token qui est développé une fois et doit donc être défini. Les autres tokens sont laissés tel quels comme avec `\noexpandarg`.

3.2 Développement des macros, argument optionnel

Les macros de ce package ne sont pas purement développables et ne peuvent donc pas être mises dans l'argument d'un `\edef`. L'imbrication des macros de ce package n'est pas permise non plus.

C'est pour cela que les macros renvoyant un résultat, c'est-à-dire toutes sauf les tests, sont dotées d'un argument optionnel venant en dernière position. Cet argument prend la forme de `[⟨nom⟩]`, où `⟨nom⟩` est une séquence de contrôle qui recevra (l'assignation se fait avec un `\edef`) le résultat de la macro, ce qui fait que `⟨nom⟩` est purement développable et peut donc se trouver dans l'argument d'un `\edef`. Dans le cas de la présence d'un argument optionnel en dernière position, aucun affichage n'aura lieu. Cela permet donc contourner les limitations évoquées dans les exemples ci dessus.

Ainsi cette construction non permise censée assigner à `\Resultat` les 4 caractères de gauche de `xstring` :

```
\edef\Resultat{\StrLeft{xstring}{4}}
```

est équivalente à :

```
\StrLeft{xstring}{4}[\Resultat]
```

Et cette imbrication non permise censée enlever le premier et le dernier caractère de `xstring` :

```
\StrGobbleLeft{\StrGobbleRight{xstring}{1}}{1}
```

se programme ainsi :

```
\StrGobbleRight{xstring}{1}[\machaine]
\StrGobbleLeft{\machaine}{1}
```

3.3 Traitement des arguments

3.3.1 Traitement à l'unité syntaxique prés

Les macros de `xstring` traitent les arguments unité syntaxique par unité syntaxique. Dans le code \TeX , une unité syntaxique¹⁵ est soit :

- une séquence de contrôle;
- un groupe, c'est à dire une suite de tokens située entre deux accolades équilibrées;
- un caractère ne faisant pas partie des 2 espèces ci dessus.

Voyons ce qu'est la notion d'unité syntaxique sur un exemple. Prenons cet argument : « `ab\textbf{xyz}cd` »

Il contient 6 unités syntaxiques qui sont : « `a` », « `b` », « `\textbf` », « `{xyz}` », « `c` » et « `d` ».

Que va t-il arriver si l'on se place sous `\noexpandarg` et que l'on demande à `xstring` de trouver la longueur de cet argument et d'en trouver le 4^e « caractère » ?

11. Pour pouvoir utiliser des lettres accentuées de façon fiable, il est nécessaire de charger le packages `\fontenc` avec l'option `[T1]`, ainsi que `\inputenc` avec l'option correspondant au codage du fichier tex.

12. Selon la règle \TeX ienne, plusieurs espaces consécutifs n'en font qu'un.

13. Attention : les accolades **doivent** être équilibrées dans les arguments !

14. C'est-à-dire qu'elle peut être mise à l'intérieur d'un `\edef`.

15. Pour les utilisateurs familiers avec la programmation \TeX , une unité syntaxique est ce qui est supprimé par la macro `\@gobble` dont le code, d'une grande simplicité, est : `\def\@gobble#1{}`

```

1 \noexpandarg
2 \StrLen{ab\textbf{xyz}cd}\par
3 \StrChar{ab\textbf{xyz}cd}{4}[\mychar]
4 \meaning\mychar

```

6
macro:->{xyz}

Il est nécessaire d'utiliser `\meaning` pour bien visualiser le véritable contenu de `\mychar` et non pas de simplement d'appeler cette séquence de contrôle, ce qui fait perdre des informations — les accolades ici. On voit qu'on n'obtient pas vraiment un « caractère », mais cela était prévisible : il s'agit d'une unité syntaxique.

3.3.2 Exploration des groupes

Par défaut, la commande `\noexploregroups` est appelée et donc dans l'argument à examiner qui contient la chaîne de tokens, `xstring` considère les groupes entre accolades comme unités syntaxiques fermées dans lesquelles `xstring` ne regarde pas.

Pour certains besoins spécifiques, il peut être nécessaire de modifier le mode de lecture des arguments et d'explorer l'intérieur des groupes entre accolades. Pour cela on peut invoquer `\exploregroups`.

Que va donner ce nouveau mode d'exploration sur l'exemple précédent ? `xstring` ne va plus compter le groupe comme une seule unité syntaxique mais va compter les unités syntaxiques se trouvant à l'intérieur, et ainsi de suite s'il y avait plusieurs niveaux d'imbrication de groupes :

```

1 \noexpandarg
2 \exploregroups
3 \StrLen{ab\textbf{xyz}cd}\par
4 \StrChar{ab\textbf{xyz}cd}{4}[\mychar]
5 \meaning\mychar

```

8
macro:->x

L'exploration des groupes peut se révéler utile pour le comptage, le calcul de position ou les tests, mais comporte une limitation lorsque l'on appelle des macros renvoyant des chaînes : lorsqu'un argument est coupé à l'intérieur d'un groupe, alors **le résultat ne tient pas compte de ce qui se trouve à l'extérieur de ce groupe**. Il faut donc utiliser ce mode en connaissance de cause lorsque l'on utilise les macros renvoyant des chaînes.

Voyons ce que cela signifie sur un exemple : mettons que l'on veuille renvoyer ce qui se trouve à droite de la 2^e occurrence de `\a` dans l'argument `\a1{\b1\a2}\a3`. Comme l'on explore les groupes, cette occurrence se trouve à l'intérieur du groupe `{\b1\a2}`. Le résultat renvoyé sera donc : `\b1`. Vérifions-le :

L'exploration des groupes¹⁶ peut ainsi changer le comportement de la plupart des macros de `xstring`, à l'exception de `\IfInteger`, `\IfDecimal`, `\IfStrEq`, `\IfEq` et `\StrCompare` qui sont insensibles au mode d'exploration en cours.

De plus, pour des raisons d'équilibrage d'accolades, 2 macros n'opèrent qu'en mode `\noexploregroups`, quelque soit le mode d'exploration en cours : `\StrBetween` et `\StrMid`.

On peut mémoriser le mode de d'exploration en cours avec `\saveexploremode`, le modifier par la suite et revenir à la situation lors de la sauvegarde avec `\restoreexploremode`.

3.4 Catcodes et macros étoilées

Les macros de `xstring` tiennent compte des catcodes des tokens constituant les arguments. Il faut donc garder à l'esprit, particulièrement lors des tests, que les tokens *et leurs catcodes* sont examinés.

Par exemple, ces 2 arguments :

$$\{\string a\string b\} \quad \text{et} \quad \{ab\}$$

ne se développent pas en 2 arguments égaux aux yeux de `xstring`. Dans le premier cas, à cause de l'emploi de la primitive `\string`, les caractères « ab » ont un catcode de 12 alors que dans l'autre cas, ils ont leurs catcodes naturels de 11. Il convient donc d'être conscient de ces subtilités lorsque l'on emploie des primitives dont les résultats sont des chaînes de caractères ayant des catcodes de 12 et 10. Ces primitives sont par exemple : `\string`, `\detokenize`, `\meaning`, `\jobname`, `\fontname`, `\romannumeral`, etc.

Pour demander aux macros de ne pas tenir compte des catcodes, on peut utiliser les macros étoilées. Après l'éventuel développement des arguments en accord avec le mode de développement, celles-ci convertissent (à l'aide d'un `\detokenize`) leur arguments en chaînes de caractères dont les catcodes sont 12 et 10 pour l'espace, avant que la macro non étoilée travaille sur ces arguments ainsi modifiés. Il faut noter que les arguments optionnels ne sont pas concernés par ces modifications et gardent leur catcode.

¹⁶ On peut consulter le fichier de test de `xstring` qui comporte de nombreux exemples et met en évidence les différences selon le mode d'exploration des groupes.

Voici un exemple :

<pre>1 \IfStrEq{\string a\string b}{ab}{vrai}{faux}\par 2 \IfStrEq*\string a\string b}{ab}{vrai}{faux}</pre>	<pre>faux vrai</pre>
--	----------------------

Les chaînes n'étant pas égales à cause des catcodes, le test est bien *négatif* dans la version non étoilée.

Attention : utiliser une macro étoilée a des conséquences ! Les arguments sont « détokénisés », il n'y a donc plus de séquence de contrôle, plus de groupements, ni aucun caractère de catcode spécial puisque tout est converti en caractères « in-offensifs » ayant le même catcode.

Ainsi, pour les macros renvoyant une chaîne, si on emploie les versions étoilées, le résultat sera une chaîne de caractères dont les catcodes sont 12, et 10 pour l'espace.

Et donc, après un « `\StrBefore*{a \b c d}{c}[\montexte]` », la séquence de contrôle `\montexte` se développera en « `a12_10\12b12_10` ».

Les macros détokénisant leur arguments par l'utilisation de l'étoile sont présentées dans la liste ci-dessous. Pour chacune d'entre elles, on peut voir en violet quels arguments seront détokénisés lorsque l'étoile sera employée :

- `\IfSubStr<[*]>[<nombre>]{<chaîne>}{<chaîneA>}{<vrai>}{<faux>}`
- `\IfSubStrBefore<[*]>[<nombre1>,<nombre2>]{<chaîne>}{<chaîneA>}{<chaîneB>}{<vrai>}{<faux>}`
- `\IfSubStrBehind<[*]>[<nombre1>,<nombre2>]{<chaîne>}{<chaîneA>}{<chaîneB>}{<vrai>}{<faux>}`
- `\IfBeginWith<[*]>{<chaîne>}{<chaîneA>}{<vrai>}{<faux>}`
- `\IfEndWith<[*]>{<chaîne>}{<chaîneA>}{<vrai>}{<faux>}`
- `\IfStrEq<[*]>{<chaîneA>}{<chaîneB>}{<vrai>}{<faux>}`
- `\IfEq<[*]>{<chaîneA>}{<chaîneB>}{<vrai>}{<faux>}`
- `\IfStrEqCase<[*]>{<chaîne>}{<chaîne1>}{<code1>}{<chaîne2>}{<code2>}`
...
{<chaîne n>}{<code n>}}[<code alternatif>]
- `\IfEqCase<[*]>{<chaîne>}{<chaîne1>}{<code1>}{<chaîne2>}{<code2>}`
...
{<chaîne n>}{<code n>}}[<code alternatif>]
- `\StrBefore<[*]>[<nombre>]{<chaîne>}{<chaîneA>}[<nom>]`
- `\StrBehind<[*]>[<nombre>]{<chaîne>}{<chaîneA>}[<nom>]`
- `\StrBetween<[*]>[<nombre1>,<nombre2>]{<chaîne>}{<chaîneA>}{<chaîneB>}[<nom>]`
- `\StrCompare<[*]>{<chaîneA>}{<chaîneB>}[<nom>]`

4 Macros avancées pour la programmation

Bien que `xstring` ait la possibilité de lire et traiter des arguments contenant du code \TeX ou \LaTeX ce qui devrait couvrir la plupart des besoins en programmation, il peut arriver pour des besoins très spécifiques que les macros décrites précédemment ne suffisent pas. Ce chapitre présente d'autres macros qui permettent d'aller plus loin ou de contourner certaines limitations.

4.1 Recherche d'un groupe, les macros `\StrFindGroup` et `\groupID`

Lorsque le mode `\exploregroups` est actif, la macro `\StrFindGroup` permet de trouver un groupe entre accolades explicites en spécifiant son identifiant :

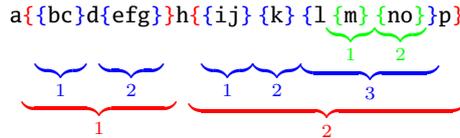
`\StrFindGroup{<argument>}{<identifiant>}[<nom>]`

Lorsque le groupe caractérisé par l'identifiant n'existe pas, une chaîne vide sera assignée à la séquence de contrôle `<nom>`. Si le groupe existe, ce groupe *avec ses accolades* sera assigné à `<nom>`.

Cet identifiant est une suite d'entiers séparés par des virgules caractérisant le groupe cherché dans l'argument. Le premier entier est le n^e groupe (d'imbrication 1) dans lequel est le groupe cherché. Puis, en se plaçant dans ce groupe, le 2^e entier est le n^e groupe dans lequel est le groupe cherché. Et ainsi de suite jusqu'à ce que l'imbrication du groupe soit atteinte.

Prenons par exemple l'argument suivant où l'on a 3 niveaux d'imbrication de groupes. Pour plus de clarté, les accolades délimitant les groupes sont colorées en rouge pour l'imbrication de niveau 1, en bleu pour le niveau 2 et en vert pour le

niveau 3. Les groupes dans chaque imbrication sont ensuite numérotés selon la règle décrite ci-dessus :



Dans cet exemple :

- le groupe `{bc}d{efg}` a donc pour identifiant **1**;
- le groupe `{ij}` a pour identifiant **2,1**;
- le groupe `{no}` a pour identifiant **2,3,2**;
- l’argument dans sa totalité « `a{bc}d{efg}h{ij}{k}{l{m}{no}}p` » a pour identifiant 0, seul cas où l’entier 0 est contenu dans l’identifiant d’un groupe.

Voici l’exemple complet :

```

1 \exploregroups
2 \expandarg
3 \def\chaine{a{bc}d{efg}h{ij}{k}{l{m}{no}}p}
4 \StrFindGroup{\chaine}{1}[\mongroupe]
5 \meaning\mongroupe\par
6 \StrFindGroup{\chaine}{2,1}[\mongroupe]
7 \meaning\mongroupe\par
8 \StrFindGroup{\chaine}{2,3,2}[\mongroupe]
9 \meaning\mongroupe\par
10 \StrFindGroup{\chaine}{2,5}[\mongroupe]
11 \meaning\mongroupe\par

```

macro:->{{bc}d{efg}}
macro:->{ij}
macro:->{no}
macro:->

Le processus inverse existe, et plusieurs macros de `xstring` donnent aussi comme information l’identifiant du groupe dans lequel elles ont fait une coupure ou trouvé une recherche. Ces macros sont : `\IfSubStr`, `\StrBefore`, `\StrBehind`, `\StrSplit`, `\StrLeft`, `\StrGobbleLeft`, `\StrRight`, `\StrGobbleRight`, `\StrChar`, `\StrPosition`.

Après l’appel à ces macros, la commande `\groupID` se développe en l’identifiant du groupe dans lequel la coupure s’est faite ou la recherche d’un argument a abouti. Lorsque la coupure ne peut avoir lieu ou que la recherche n’a pas abouti, `\groupID` est vide. Évidemment, l’utilisation de `\groupID` n’a de sens que lorsque le mode `\exploregroups` est actif, et quand les macros ne sont pas étoilées.

Voici quelques exemples avec la macro `\StrChar` :

```

1 \exploregroups
2 char 1 = \StrChar{a{b{cd}{e{f}g}h}ijkl}{1}\qqquad
3 \string\groupID = \groupID\par
4 char 4 = \StrChar{a{b{cd}{e{f}g}h}ijkl}{4}\qqquad
5 \string\groupID = \groupID\par
6 char 6 = \StrChar{a{b{cd}{e{f}g}h}ijkl}{6}\qqquad
7 \string\groupID = \groupID\par
8 char 20 = \StrChar{a{b{cd}{e{f}g}h}ijkl}{20}\qqquad
9 \string\groupID = \groupID

```

char 1 = a \groupID= 0
char 4 = d \groupID= 1,1
char 6 = f \groupID= 1,1,2,1
char 20 = \groupID=

4.2 Coupure d’une chaîne, la macro `\StrSplit`

Voici la syntaxe de cette macro :

`\StrSplit{<chaîne>}{<nombre>}{<chaîneA>}{<chaîneB>}`

La `<chaîne>`, est coupée en deux chaînes juste après l’unité syntaxique se situant à la position `<nombre>`. La partie gauche est assigné à la séquence de contrôle `<chaîneA>` et la partie droite à `<chaîneB>`.

Cette macro renvoie *deux chaînes* et donc **n’affiche rien**. Par conséquent, elle ne dispose pas de l’argument optionnel en dernière position.

- ▷ Si `<nombre> ≤ 0`, `<chaîneA>` sera vide et `<chaîneB>` contiendra la totalité de `<chaîne>`;
- ▷ Si `<nombre> ≥ <longueurChaîne>`, `<chaîneA>` contiendra la totalité de `<chaîne>` et `<chaîneB>` sera vide;
- ▷ Si `<chaîne>` est vide `<chaîneA>` et `<chaîneB>` seront vides, quelque soit l’entier `<nombre>`.

<pre> 1 \def\seprouge{{\color{red}} }} 2 \StrSplit{abcdef}{4}{\csA}{\csB} \csA\seprouge\csB \par 3 \StrSplit{a b c}{2}{\csA}{\csB} \csA\seprouge\csB \par 4 \StrSplit{abcdef}{1}{\csA}{\csB} \csA\seprouge\csB \par 5 \StrSplit{abcdef}{5}{\csA}{\csB} \csA\seprouge\csB \par 6 \StrSplit{abcdef}{9}{\csA}{\csB} \csA\seprouge\csB \par 7 \StrSplit{abcdef}{-3}{\csA}{\csB} \csA\seprouge\csB </pre>	<pre> abcd ef a b c a bcdef abcde f abcdef abcdef </pre>
---	---

Lorsque l'exploration des groupes est activée, et que l'on demande une coupure en fin de groupe, alors une chaîne contiendra la totalité du groupe tansque l'autre sera vide comme on le voit sur cet exemple :

<pre> 1 \exploregroups 2 \StrSplit{ab{cd{ef}gh}ij}{6}\strA\strB 3 \meaning\strA\par 4 \meaning\strB </pre>	<pre> macro:->ef macro:-> </pre>
--	--

Une version étoilée de cette macro existe : dans ce cas, la coupure se fait juste avant la prochaine unité syntaxique qui suit l'unité syntaxique désirée. La version étoilée ne donne des résultats différents de la version normale que lorsque la n^e unité syntaxique est à la fin d'un groupe auquel cas, la coupure intervient non pas après cette unité syntaxique mais *avant* la prochaine unité syntaxique, que `\StrSplit` atteint en fermant autant de groupes que nécessaire.

<pre> 1 \exploregroups 2 Utilisation sans \'etoile : \par 3 \StrSplit{ab{cd{ef}gh}ij}{6}\strA\strB 4 \meaning\strA\par 5 \meaning\strB\par 6 \string\groupID\ = \groupID\par\medskip 7 Utilisation avec \'etoile : \par 8 \StrSplit*{ab{cd{ef}gh}ij}{6}\strA\strB 9 \meaning\strA\par 10 \meaning\strB\par 11 \string\groupID\ = \groupID </pre>	<pre> Utilisation sans étoile : macro:->ef macro:-> \groupID = 1,1 Utilisation avec étoile : macro:->cd{ef} macro:->gh \groupID = 1,1 </pre>
--	---

4.3 Assigner un contenu verb, la macro `\verbtocs`

La macro `\verbtocs` permet de lire le contenu d'un « verb » qui peut contenir tous les caractères spéciaux : &, ~, \, {, }, ..., #, \$, ^ et %. Les caractères « normaux » gardent leur catcodes naturels, sauf les caractères spéciaux qui prennent un catcode de 12. Ensuite, ces caractères sont assignés à une séquence de contrôle. La syntaxe complète est :

`\verbtocs{<nom>|<caractères>|}`

`<nom>` est le nom d'une séquence de contrôle qui recevra à l'aide d'un `\edef` les `<caractères>`. `<nom>` contiendra donc des caractères de catcodes 12 ou 10 pour l'espace.

Par défaut, le token délimitant le contenu verb est « | », étant entendu que ce token ne peut être à la fois le délimiteur et être contenu dans ce qu'il délimite. Au cas où on voudrait lire un contenu verb contenant « | », on peut changer à tout moment le token délimitant le contenu verb par la macro :

`\setverbdelim{<token>}`

Tout `<token>` de catcode 12 peut être utilisé¹⁷.

Concernant ces arguments verb, il faut tenir compte des deux points suivants :

- tous les caractères se trouvant avant `|<caractères>|` seront ignorés ;
- à l'intérieur des délimiteurs, tous les espaces sont comptabilisés même s'ils sont consécutifs.

Exemple :

<pre> 1 \verbtocs{resultat} a & b{ c% d\$ e \f 2 J'affiche le r^esultat : \par\resultat </pre>	<pre> J'affiche le résultat : a & b{ c% d\$ e \f </pre>
--	---

¹⁷. Plusieurs tokens peuvent être utilisés au risque d'alourdir la syntaxe de `\verbtocs` ! Pour cette raison, un avertissement sera émis si l'argument de `\setverbdelim` contient plusieurs tokens.

4.4 Tokenisation d'un texte vers une séquence de contrôle, la macro `\tokenize`

Le processus inverse de ce qui a été vu au dessus consiste à interpréter une suite de caractères en tokens. Pour cela, on dispose de la macro :

$$\backslash\text{tokenize}\{\langle\text{nom}\rangle\}\{\langle\text{texte}\rangle\}$$

$\langle\text{texte}\rangle$ est développé le plus possible si l'on a invoqué `\fullexpandarg` ; il n'est pas développé si l'on a invoqué `\noexpandarg` ou `\expandarg`. Après développement éventuel, le $\langle\text{texte}\rangle$ est transformé en tokens puis assigné à l'aide d'un `\def` à la séquence de contrôle $\langle\text{nom}\rangle$.

Voici un exemple où l'on détokenise un argument, on affiche le texte obtenu, puis on transforme ce texte en ce que l'argument était au début ; enfin, on affiche le résultat de la tokenisation :

<pre> 1 \verbtocs{\text} \textbf{a} \$\frac{1}{2}\$ 2 texte : \text 3 \tokenize{\resultat}{\text}\par 4 r'esultat : \resultat </pre>	<pre> texte : \textbf{a} \$\frac{1}{2}\$ résultat : a $\frac{1}{2}$ </pre>
--	---

Il est bien évident à la dernière ligne, que l'appel à la séquence de contrôle `\resultat` est ici possible puisque les séquences de contrôle qu'elle contient sont définies.

4.5 Développement contrôlé, les macros `\StrExpand` et `\scans`

La macro `\StrExpand` développe les tokens de la chaîne passée en argument. Voici sa syntaxe :

$$\backslash\text{StrExpand}[\langle\text{nombre}\rangle][\langle\text{chaîne}\rangle]\{\langle\text{nom}\rangle\}$$

Le $\langle\text{nombre}\rangle$ vaut 1 par défaut et représente le nombre de développement(s) que doit subir la $\langle\text{chaîne}\rangle$ de tokens. Le $\langle\text{nom}\rangle$ est le nom d'une séquence de contrôle à laquelle est assigné le résultat, une fois que tous les tokens aient été développé le nombre de fois demandé.

La macro opère séquentiellement et par passe : chaque token est remplacé par son 1-développement, et le token suivant subit le même traitement jusqu'à ce que la chaîne ait été parcourue. Ensuite, s'il reste des niveaux de développement à faire, une nouvelle passe est initiée, et ainsi de suite jusqu'à ce que le nombre de développements voulu aient été exécutés.

Voici un exemple :

<pre> 1 \def\csA{1 2} 2 \def\csB{a \csA} 3 \def\csC{\csB\space} 4 \def\csD{x\csA y}\csB{\csC z} 5 D'veveloppement de \string\csD\ au\par 6 \StrExpand[0]{\csD}{\csE} niveau 0 : 7 \detokenize\expandafter{\csE}\par 8 \StrExpand[1]{\csD}{\csE} niveau 1 : 9 \detokenize\expandafter{\csE}\par 10 \StrExpand[2]{\csD}{\csE} niveau 2 : 11 \detokenize\expandafter{\csE}\par 12 \StrExpand[3]{\csD}{\csE} niveau 3 : 13 \detokenize\expandafter{\csE}\par 14 \StrExpand[4]{\csD}{\csE} niveau 4 : 15 \detokenize\expandafter{\csE} </pre>	<pre> Développement de \csD au niveau 0 : \csD niveau 1 : x\csA y}\csB {\csC z} niveau 2 : x{1 2y}a \csA {\csB \space z} niveau 3 : x{1 2y}a 1 2{a \csA z} niveau 4 : x{1 2y}a 1 2{a 1 2 z} </pre>
--	--

La macro agit séquentiellement et chaque token est développé isolément sans tenir compte de ce qui suit. On ne peut donc développer que des tokens qui se suffisent à eux même et dont le développement ne nécessite aucun autre token. Ainsi, « `\iftrue A\else B\fi` », bien que se développant en « A » ne peut être mis dans l'argument de `\StrExpand`, et l'instruction :

$$\backslash\text{StrExpand}\{\text{iftrue } A\text{else } B\text{fi}\}\backslash\text{resultat}$$

fera échouer la compilation puisque le premier token « `\iftrue` » sera développé *seul*, c'est-à-dire sans son `\fi` correspondant, ce qui fâchera $\text{T}_\text{E}\text{X}$!

Les règles habituelles de lecture des arguments sont en vigueur, à savoir qu'un espace suivant une séquence de contrôle est ignoré, et plusieurs espaces consécutifs n'en font qu'un. Ces règles ne s'appliquent pas pour des espaces explicitement demandés avec `\space` ou `_`¹⁸.

Le développement de ce qui se trouve à l'intérieur des groupes est *indépendant* du mode d'exploration des groupes : cette macro possède ses propres commandes pour développer ce qui est dans les groupes ou pas. Par défaut, les tokens se

¹⁸. À ce propos, `\space` n'a pas la même signification que `_`. La première séquence de contrôle se *développe* en un espace tandis que la deuxième est une primitive $\text{T}_\text{E}\text{X}$ qui *affiche* une espace. Comme toutes les primitives, cette dernière est son propre développement.

trouvant à l'intérieur des groupes sont développés, mais on peut demander à ce que ce développement ne se fasse pas en invoquant `\noexpandingroups` et revenir au comportement par défaut par `\expandingroups`.

On peut détokeriser le résultat obtenu par la macro `\StrExpand` avec la macro `\scancs` dont la syntaxe est :

```
\scancs[⟨nombre⟩]{⟨nom⟩}{⟨chaîne⟩}
```

Le `⟨nombre⟩` vaut 1 par défaut et représente le nombre de développement(s) que doit subir chaque token constituant la `⟨chaîne⟩`. Le `⟨nom⟩` est le nom d'une séquence de contrôle à laquelle est assigné le résultat, une fois que tous les tokens aient été développés le nombre de fois demandé et ensuite détokerisés.

`\scancs` a été conservée pour garder une compatibilité avec des précédentes versions de `xstring`. Pour les mêmes raisons, sa syntaxe n'est pas cohérente avec la syntaxe des autres macros. Cette macro, devenue triviale, prend le résultat de `\StrExpand` et lui applique simplement un `\detokenize`.

4.6 À l'intérieur d'une définition de macro

Avec le verbatim, certaines difficultés surviennent lorsque l'on se trouve à l'intérieur de la définition d'une macro, c'est-à-dire entre les accolades suivant un `\def\macro` ou un `\newcommand\macro`.

Pour les mêmes raison qu'il est interdit d'employer la commande `\verb` à l'intérieur de la définition d'une macro, les arguments `verb` du type `|⟨caractères⟩|` sont également interdits, ce qui disqualifie la macro `\verbtocs`. Il faut donc observer la règle suivante :

Ne pas utiliser la macro `\verbtocs` à l'intérieur de la définition d'une macro.

Mais alors, comment faire pour manipuler des arguments textuels `verb` et « verbatimiser » dans les définitions de macro ?

Il y a la primitive `\detokenize` de ε -TeX, mais elle comporte des restrictions, entre autres :

- les accolades doivent être équilibrées ;
- les espaces consécutifs sont ignorés ;
- les signes % sont interdits ;
- une espace est ajoutée après chaque séquence de contrôle.

Il est préférable d'utiliser la macro `\scancs`, et définir avec `\verbtocs` à l'extérieur des définitions de macros, des séquences de contrôle contenant des caractères spéciaux détokerisés. On pourra aussi utiliser la macro `\tokenize` pour transformer le résultat final (qui est une chaîne de caractères) en une séquence de contrôle. On peut voir des exemples utilisant ces macros page 20, à la fin de ce manuel.

Dans l'exemple artificiel¹⁹ qui suit, on écrit une macro qui met son argument entre accolades. Pour cela, on définit en dehors de la définition de la macro 2 séquences de contrôles `\Ob` et `\Cb` contenant une accolade ouvrante et une accolade fermante de catcodes 12. Ces séquences de contrôle sont ensuite développées et utilisées à l'intérieur de la macro pour obtenir le résultat voulu :

```

1 \verbtocs{\Ob}{|{|
2 \verbtocs{\Cb}{|}|
3 \newcommand\bracearg[1]{%
4   \def\text{#1}%
5   \scancs{\result}{\Ob\text\Cb}%
6   \result}
7
8 \bracearg{xstring}\par
9 \bracearg{a}

```

{xstring}
{a }

4.7 La macro `\StrRemoveBraces`

Pour des utilisations spéciales, on peut désirer retirer les accolades délimitant les groupes dans un argument.

On peut utiliser la macro `\StrRemoveBraces` dont voici la syntaxe :

```
\StrRemoveBraces{⟨chaîne⟩}[⟨nom⟩]
```

Cette macro est sensible au mode d'exploration, et retirera *toutes* les accolades avec `\exploregroups` alors qu'elle ne retirera que les accolades des groupes de plus bas niveau avec `\noexploregroups`.

¹⁹. On peut agir beaucoup plus simplement en utilisant la commande `\detokenize`. Il suffit de définir la macro ainsi :
`\newcommand\bracearg[1]{\detokenize{#1}}`

```

1 \noexploregroups
2 \StrRemoveBraces{a{b{c}d}e{f}g}[\mycs]
3 \meaning\mycs
4
5 \exploregroups
6 \StrRemoveBraces{a{b{c}d}e{f}g}[\mycs]
7 \meaning\mycs

```

macro:->ab{c}defg
macro:->abcdefg

4.8 Exemples d'utilisation en programmation

Voici quelques exemples très simples d'utilisation des macros comme on pourrait en rencontrer en programmation.

4.8.1 Exemple 1

On cherche à remplacer les deux premiers `\textit` par `\textbf` dans la séquence de contrôle `\myCS` qui contient :

```
\textit{A}\textit{B}\textit{C}
```

On cherche évidemment à obtenir `\textbf{A}\textbf{B}\textit{C}` qui affiche : **ABC**

Pour cela, on va développer les arguments des macros une fois avant qu'elles les traitent, en invoquant la commande `\expandarg`.

Ensuite, on définit `\pattern` qui est le motif à remplacer, et `\replace` qui est le motif de substitution. On travaille token par token puisque `\expandarg` a été appelé, il suffit d'invoquer `\StrSubstitute` pour faire les 2 substitutions.

```

1 \expandarg
2 \def\myCS{\textit{A}\textit{B}\textit{C}}
3 \def\pattern{\textit}
4 \def\replace{\textbf}
5 \StrSubstitute[2]{\myCS}{\pattern}{\replace}

```

ABC

Pour éviter de définir les séquences de contrôle `\pattern` et `\replace`, on aurait pu utiliser un leurre comme par exemple une séquence de contrôle qui se développe en « rien » comme `\empty`, et coder de cette façon :

```
\StrSubstitute[2]{\myCS}{\empty\textit}{\empty\textbf}
```

Ainsi, `\empty` est développée en « rien » et il reste dans les 2 derniers arguments les séquences de contrôles significatives `\textit` et `\textbf`.

La séquence de contrôle `\empty` est donc un « hack » pour `\expandarg` : elle permet de bloquer le développement du 1^{er} token ! On aurait d'ailleurs pu utiliser `\noexpand` au lieu de `\empty` pour obtenir le même résultat.

4.8.2 Exemple 2

On cherche ici à écrire une commande qui efface n unités syntaxiques dans une chaîne à partir d'une position donnée, et affecte le résultat dans une séquence de contrôle dont on peut choisir le nom.

On va appeler cette macro `StringDel` et lui donner la syntaxe :

```
\StringDel{chaîne}{position}{n}{\nom_resultat}
```

On peut procéder ainsi : sauvegarder la chaîne se trouvant juste avant la position. Ensuite enlever $n + \text{position}$ unités syntaxiques à la chaîne initiale, et concaténer ce résultat à ce qui a été sauvegardé auparavant. Cela donne le code suivant :

```

1 \newcommand\StringDel[4]{%
2   \begingroup
3   \expandarg% portee locale au groupe
4   \StrLeft{\empty#1}{\number\numexpr#2-1}[#4]%
5   \StrGobbleLeft{\empty#1}{\numexpr#2+#3-1}[\StrA]%
6   \expandafter\expandafter\expandafter\endgroup
7   \expandafter\expandafter\expandafter\def
8   \expandafter\expandafter\expandafter#4%
9   \expandafter\expandafter\expandafter
10  {\expandafter#4\StrA}%
11 }
12
13 \noexploregroups
14 \StringDel{abcdefgh}{2}{3}{\cmd}
15 \meaning\cmd
16
17 \StringDel{a\textbf{1}b\textbf{2c}3d}{3}{4}{\cmd}
18 \meaning\cmd

```

macro:->aefgh
macro:->a\textbf 3d

Pour la concaténation, on aurait pu procéder différemment en utilisant la macro `\g@addto@macro` de \TeX . Cela évite aussi ces laborieux « ponts » d'`\expandafter`. Il suffit alors de remplacer l'assignation et la sortie du groupe se trouvant entre les lignes 6 à 10 par ²⁰ :

```
\expandafter\g@addto@macro\expandafter#4\expandafter{\StrA}\endgroup
```

4.8.3 Exemple 3

Cherchons à écrire une macro `\tofrac` qui transforme une écriture du type « a/b » par « $\frac{a}{b}$ ».

Tout d'abord, annulons le développement des arguments avec `\noexpandarg` : nous n'avons pas besoin de développement ici. Il suffit d'isoler ce qui se trouve avant et après la 1^{re} occurrence de « / » (on suppose qu'il n'y a qu'une seule occurrence), le mettre dans les séquences de contrôle `\num` et `\den` et simplement appeler la macro \TeX `\frac` :

```
1 \noexpandarg
2 \newcommand\tofrac[1]{%
3   \StrBefore{#1}{/}[\num]%
4   \StrBehind{#1}{/}[\den]%
5   $\frac{\num}{\den}$%
6 }
7 \tofrac{15/9}
8 \tofrac{u_{n+1}/u_n}
9 \tofrac{a^m/a^n}
10 \tofrac{x+\sqrt{x}}{\sqrt{x^2+x+1}}
```

$$\frac{15}{9} \frac{u_{n+1}}{u_n} \frac{a^m}{a^n} \frac{x+\sqrt{x}}{\sqrt{x^2+x+1}}$$

4.8.4 Exemple 4

Soit une phrase composée de texte. Dans cette phrase, essayons construire une macro qui mette en gras le 1^{er} mot qui suit un mot donné. On entend par mot une suite de caractère ne commençant ni ne finissant par un espace. Si le mot n'existe pas dans la phrase, rien n'est fait.

On va écrire une macro `\grasapres` qui effectue ce travail. On appelle `\StrBehind` qui assigne à `\mot` ce qui se trouve après la 1^{re} occurrence du mot (précédé et suivi de son espace). Ensuite, le mot à mettre en gras est ce qui se trouve avant le 1^{er} espace dans la séquence de contrôle `\mot`. Remarquons que ceci reste vrai même si le mot à mettre en gras est le dernier de l'argument car un espace a été rajouté à la fin de l'argument par `{#1 }` lors de l'appel à `\StrBehind`. Remarquons aussi que `\expandarg` a été appelé et donc, le premier token de l'argument `\textbf{\mot}` est développé 1 fois, *lui aussi!* Cela est possible (heureusement sinon, il aurait fallu faire autrement et utiliser le hack de l'exemple précédent) puisque le 1-développement de cette macro de \TeX est « `\protect\textbf` » ²¹.

```
1 \newcommand\grasapres[2]{%
2   \noexpandarg
3   \StrBehind[1]{#1 }{ #2 }[\mot]%
4   \expandarg
5   \StrBefore{\mot}{ }[\mot]%
6   \StrSubstitute[1]{#1}{\mot}{\textbf{\mot}}%
7 }
8
9 \grasapres{Le package xstring est nouveau}{package}
10 \grasapres{Le package xstring est nouveau}{ckage}
11 \grasapres{Le package xstring est nouveau}{est}
```

Le package **xstring** est nouveau
Le package xstring est nouveau
Le package xstring est **nouveau**

4.8.5 Exemple 5

Soit un argument commençant par au moins 3 séquences de contrôles avec leurs éventuels arguments. Comment intervertir les 2 premières séquences de contrôle de telle sorte qu'elles gardent leurs arguments ? On va pour cela écrire une macro `\swaptwofirst`.

Cette fois ci, on ne peut pas chercher le seul caractère « \ » (de catcode 0) dans un argument. Nous serons obligé de détokeriser l'argument, c'est ce que fait `\scans[0]\chaîne{#1}` qui met le résultat dans `\chaîne`. Ensuite, on cherchera dans cette séquence de contrôle les occurrences de `\antislash` qui contient le caractère « \ » de catcode 12, assigné avec

20. À condition d'avoir provisoirement changé le code de catégorie de « @ » en écrivant la macro entre les commandes `\makeatletter` et `\makeatother`
21. En toute rigueur, il aurait fallu écrire :

```
\StrSubstitute[1]{#1}{\mot}{\expandafter\textbf\expandafter{\mot}}
```

De cette façon, dans `{\expandafter\textbf\expandafter{\mot}}`, la séquence de contrôle `\mot` est développée **avant** que l'appel à la macro ne se fasse. Cela est dû à `\expandafter` placé en début d'argument qui est développé à cause de `\expandarg` et grâce à l'autre `\expandafter`, provoque le développement de `\mot`

un `\verbtocs` écrit *en dehors*²² du corps de la macro. La macro se termine par une retokenisation, une fois que les chaînes `\avant` et `\apres` aient été échangées.

```

1 \verbtocs{\antislash}|\|
2 \newcommand\swaptwofirst[1]{%
3   \begingroup
4   \fullellarg
5   \scans[0]\chaîne{#1}%
6   \StrBefore[3]{\chaîne}{\antislash}[\firsttwo]%
7   \StrBehind{\chaîne}{\firsttwo}[\others]
8   \StrBefore[2]{\firsttwo}{\antislash}[\avant]
9   \StrBehind{\firsttwo}{\avant}[\apres]%
10  \tokenize\myCS{\apres\avant\others}%
11  \myCS
12  \endgroup
13 }
14
15 \swaptwofirst{\underline{A}\textbf{B}\textit{C}}
16
17 \swaptwofirst{\Large\underline{A}\textbf{B}123}

```

BAC
AB123

4.8.6 Exemple 6

Dans une chaîne, on cherche ici à isoler le n^e mot se trouvant entre 2 délimiteurs précis. Pour cela on écrira une macro `\findword` admettant comme argument optionnel le délimiteur de mot (l'espace par défaut), 1 argument contenant la chaîne, 1 argument contenant le nombre correspondant au n^e mot cheché.

La macro `\findword` utilise `\StrBetween` et `\numexpr` de façon astucieuse, et profite du fait que `\StrBetween` n'explore pas les groupes :

```

1 \newcommand\findword[3][ ]{%
2   \StrBetween[#3,\numexpr#3+1]{#1#2#1}{#1}{#1}%
3 }
4 \noexpandarg
5 |\findword{a bc d\textbf{e f} gh}{3}|
6
7 |\findword[\nil]{1 \nil 2 3 \nil4\nil5}{2}|

```

|de f|
|2 3|

On peut observer que le résultat de la ligne 7 qui est « 2_3_ » ne commence *pas* par une espace puisque dans le code, cet espace suit une séquence de contrôle — ici `\nil`.

*
* *

C'est tout, j'espère que ce package vous sera utile!
Merci de me signaler par [email](#) tout bug ou toute proposition d'amélioration...

Christian TELLECHEA

²². En effet, la macro `\verbtocs` et son argument `verb` est interdite à l'intérieur de la définition d'une macro.