

Chapitre 5 Pointeurs et chaînes

Marco Lavoie

Langage C++ 14278 ORD



Dans ce chapitre nous discutons de l'une des caractéristiques les plus puissantes et le plus difficiles à maîtriser du langage C++ : le pointeur. La manipulation de pointeurs est ce qui distingue principalement le C++ des autres langages tels que Java ou C#. En effet, ces langages ne supportent pas la manipulation de pointeurs, caractéristique essentielle au développement d'applications embarqués (tels que la programmation d'appareils électroménagers, de véhicules motorisés ou d'appareils médicaux). C'est pourquoi le langage C++ est si populaire dans l'industrie du développement d'applications industrielles et commerciales.



Objectifs

- Utiliser les pointeurs
 - Pour passer des arguments à des fonctions en appel par référence via pointeurs
- Comprendre les relations entre les pointeurs, les tableaux et les chaînes
 - Déclarer et utiliser des tableaux pour chaînes de caractères

Marco Lavoie

14728 ORD - Langage C++

Comme nous l'avons vu au chapitre 3, nous pouvons utiliser des références (opérateur &) pour effectuer des appels par référence. Les pointeurs permettent aux programmes de simuler ce genre d'appels et de créer et manipuler des structures de données dynamiques ou, plus précisément, des structures de données capables de grossir et de rétrécir comme les listes chaînées, les queue, les piles et les arborescences (techniques de programmation que vous étudierez aux étapes ultérieures de TGI).



🔼 Aperçu

- Déclaration de variables pointeurs
- · Opérateurs de pointeurs
- Appels de fonctions par référence via pointeurs
- Arithmétique des pointeurs
- · Relation entre les pointeurs et les tableaux
 - Tableaux de pointeurs
- · Introduction au traitement de caractères et de chaînes
 - Principes de base
 - Bibliothèque de manipulation de chaînes de caractères

Marco Lavoie

14728 ORD - Langage C++

Dans ce chapitre, nous expliquons et analysons les concepts de base des pointeurs ainsi que les relations intimes entre les tableaux, les pointeurs et les chaînes de caractères.

Nous verrons l'emploi des pointeurs avec les structures au chapitre 6, tandis qu'aux chapitres 9 et 10 nous apprendrons que la programmation orientée objets s'effectue au moyen de pointeurs et de références.



Mark Introduction

- Les pointeurs : une caractéristique
 - parmi les plus puissantes en programmation
 - parmi les plus difficiles à maîtriser
- Pas unique au C++
 - C, Pascal, assembleur, etc.
- Plusieurs cours de TGI exploitent les pointeurs
 - Classes et structures de données I et II
 - Programmation assembleur
 - Systèmes embarquées I et II
 - Protocoles de communication
 - Algorithmes d'analyse syntaxique

14728 ORD - Langage C++

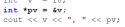
Comme mentionné précédemment, les pointeurs sont l'une des caractéristiques les plus puissantes et le plus difficiles à maîtriser du langage C++. Plusieurs langages tels que Java et C# exploitent intrinsèquement les pointeurs mais de façon cachée au programmeur, principalement car la gestion des pointeurs dans un programme est particulièrement difficile et occasionne souvent des bogues dans ceux-ci. Conséquemment ces langages de programmation « cachent » le concept de pointeurs aux programmeurs afin de lui éviter ces difficultés. Par contre, le langage C++ dévoile les pointeurs aux programmeurs, laissant à ces derniers le loisir de les manipuler dans leurs programmes. À cette liberté est cependant rattaché un niveau de difficulté accru dans la programmation d'applications complexes.

Hérité du langage C (l'ancêtre du C++), les pointeurs seront exploités intensivement dans d'autres cours de votre programme d'études, aux étapes 3 à 6.



🔼 déréférencement

- Les pointeurs implantent le principe de déréférencement
 - Une variable contient une valeur
 - Un pointeur est une variable contenant l'adresse d'une variable contenant une valeur



• pv contient l'adresse de la variable v



18, 0012FF60

14728 ORD - Langage C++

Les variables pointeurs contiennent des adresses générale, une En règle conventionnelle contient directement une valeur spécifique (p.ex. la variable v contient la valeur 18), tandis qu'une variable pointeur contient plutôt l'adresse d'une variable qui, elle, contient une valeur spécifique (p.ex. pv contient l'adresse de v qui elle contient 18). Ainsi, un nom de variable référence directement une valeur, alors qu'un pointeur référence indirectement une valeur. Référencer une valeur via un pointeur s'appelle la déréférencement.

L'opérateur d'adresse & retourne l'adresse de la variable spécifiée (p.ex. &v retourne l'adresse du premier octet de mémoire alloué à v). Dans l'exemple ci-contre, cette adresse (i.e. adresse mémoire de v obtenue par &v) est stockée dans la variable pointeur pv. On dit que le pointeur pv « pointe » à v, tel qu'illustré par la flèche ci-



🔼 Opérateur *

- Permet de
 - Déclarer une variable pointeur
 - Déréférencer au contenu de la variable pointée

```
int v = 18;
int *pv = &v;
                " << pv << ", " << *pv << endl;
*pv = 33;
cout << v << ", " << pv << ", " << *pv << endl;
```



Marco Lavoie

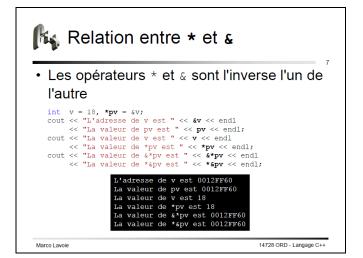
14728 ORD - Langage C++

L'opérateur de déréférencement * a deux fonctions :

- 1. Il sert à déclarer des variables pointeurs.
- 2. Il est à déréférencer un pointeur afin d'accéder à l'espace mémoire pointé.

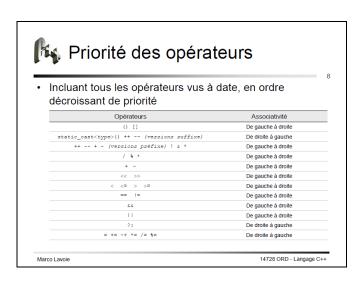
Dans l'exemple ci-contre, l'opérateur * est utilisé une première fois pour déclarer la variable pointeur pv : pv est une variable pointeur (int *pv;) destinée à stockée l'adresse d'une valeur entière (int *pv;).

La seconde utilisation de l'opérateur * exploite la déréférence pour modifié le contenu du bloc mémoire alloué à la variable v via l'adresse de ce bloc stockée dans pv. Ainsi, les deux affectations v = 33 et *pv = 33 sont équivalentes puisqu'elles font référence au même bloc mémoire.



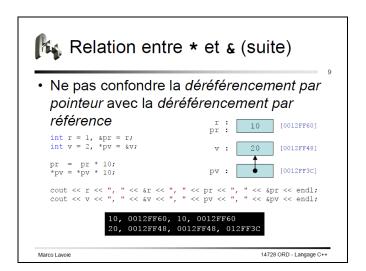
Il y a une relation intrinsèque entre l'opérateur d'adresse \mathbf{c} et celui de déréférencement \mathbf{v} : l'opérateur d'adresse sert à obtenir l'adresse d'un bloc mémoire (tel que celui alloué par le compilateur à la variable \mathbf{v}) et l'opérateur de déréférencement sert à référencer le contenu d'un bloc mémoire via son adresse (tel que l'accès au contenu de la variable \mathbf{v} via son adresse).

Observez l'affichage produit par les deux dernières instructions de l'exemple ci-contre : les opérateurs & et * sont l'inverse l'un de l'autre puisque le même résultat (l'adresse stockée dans pv) apparaît lorsqu'on les applique tous deux consécutivement à pv.



Le tableau ci-contre illustre le niveau de priorité et l'associativité des opérateurs vus à date. Attention à ne pas confondre l'opérateur de déréférencement * avec l'opérateur arithmétique de multiplication *, qui exploitent le même symbole (l'étoile) mais ont un niveau de priorité distinct.

Il est important de noter que l'opérateur de déréférencement dispose d'un niveau de priorité plus élevé que celui des opérateurs arithmétiques. Ainsi, l'expression *pv + 1 est équivalente à (*pv) + 1, et non *(pv + 1). Cette distinction sera importante lorsque nous verrons plus loin l'arithmétique des pointeurs.



Les programmeurs débutants en C++ confondent parfois les *variables alias* (déclarés avec l'opérateur d'adresse s) et les *variables pointeurs* (déclarées avec l'opérateur de déréférencement *), puisque toutes deux servent généralement à référencer le contenu d'une autre variable.

Dans l'exemple ci-contre, la variable alias **pr** et la variable pointeur **pv** référencent toutes deux le contenu d'une autre variable (**pr** référence le contenu de **r** et **pv** référence le contenu de **v**). Cependant la variable alias **pr** ne nécessite aucun autre opérateur pour accéder au contenu de **r**, alors que la variable pointeur **pv** requiert l'opérateur de déréférencement * pour accéder au contenu de **v**.

La variable alias **pr** ne requiert aucun autre opérateur pour accéder au contenu de **r** car, en fait, **pr** et **r** font référence au même bloc mémoire (adresse **0012FF60**) alors que **v** et **pv** sont deux variables distinctes, telle qu'illustrée dans l'affichage ci-contre.



Paramètres pointeurs

- Alternative aux paramètres de référence
 - Permet à une fonction de modifier le contenu des arguments d'invocation
- Rappel : paramètres de référence

```
void swap ( int &x , int &y ) {
 int temp = x;
x = y;
y = temp;
                                                  std::cout << setw( 3 );
                                                   std::cout << a << b << std::endl;
                                                  swap( a, b );
std::cout << a << b << std::endl;</pre>
```

En C++, il existe trois façons de passer des arguments vers une fonction lors de l'invocation de celle-ci :

- 1. l'appel par valeur,
- 2. l'appel par référence via référence, et
- 3. l'appel par référence via pointeur.

Dans le chapitre trois nous avons présenté les deux premières façons.

L'exemple ci-contre illustre le passage d'arguments par appel par référence via référence : l'opérateur d'adresse & est utilisé pour déclarer les paramètres de la fonction swap comme étant des alias des arguments fournis lors de l'invocation de swap dans main. Ainsi, lorsque swap manipule le contenu des paramètres x et y, elle manipule en réalité le contenu des variables a et b de main.



🄼 Paramètres pointeurs (suite)

- Paramètre pointeur : reçoit l'adresse de l'argument correspondant
 - Permet ainsi de modifier le contenu de cet argument

```
// Transmutation des arguments
void swap( int *x, int *y ) {
  int temp = *x;
                                                                     int main() {
  int a = 10, b = 20;
                                                                        std::cout << setw(3);
                                                                         std::cout << a << b << std::endl;
swap( &a, &b );
std::cout << a << b << std::endl;</pre>
                                                                        return 0;
```

Marco Lavoi

14728 ORD - Langage C++

14728 ORD - Langage C++

L'appel par référence via pointeur exploite l'opérateur d'adresse & conjointement avec l'opérateur de déréférencement * pour faire en sorte que la fonction manipule le contenu des arguments. La version de la fonction swap présentée ci-contre déclare paramètres comme étant des pointeurs, c.à.d. des variables destinées à contenir l'adresse de d'autres variables. Conséquemment, les instructions du corps de l'opérateur fonction swap exploitent déréférencement * pour accéder au contenu des arguments a et b par déréférencement via leurs adresses respectives stockées dans x et y. De même, l'invocation dans main transmet à swap les adresses des variables a et b, qui seront respectivement affectées aux paramètres x et y.

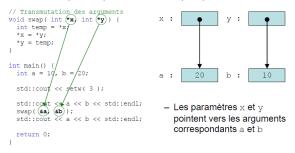
La page suivante de la présentation illustre bien cette relation entre les paramètres pointeurs de swap et les arguments fournis par main lors de l'invocation.



Marco Lavoie

🄼 Paramètres pointeurs (suite)

Exécution pas à pas de l'exemple précédent



Puisque le paramètre pointeur x contient l'adresse de l'argument a, et le paramètre pointeur y contient l'adresse de l'argument b, les instructions de swap manipulent les valeurs stockées dans a et b via ses paramètres pointeurs et l'opérateur de déréférencement *. Ainsi, l'instruction *x = *y; dans swap extrait indirectement la valeur de b (y contenant l'adresse de b, *y faisant référence au bloc de mémoire alloué à b) pour l'attribuer à la variable a (x contient l'adresse de a, donc *x fait référence au bloc de mémoire alloué à a).

L'opérateur de déréférencement * peut être vu comme signifiant « suivre la flèche » dans l'illustration ci-contre. Ainsi, dans l'instruction suivante :

```
cout << x << *x;
```

La première valeur affichée est le contenu de x (soit l'adresse de a), alors que la seconde valeur affichée est 20, soit le contenu de la variable a, accédée indirectement via son adresse stockée dans x.

© Marco Lavoie Page 5.4

14728 ORD - Langage C++



🄼 Paramètres pointeurs vs références

· Même fonctionnalité, mais attention aux distinctions dans le code source

```
Paramètres pointeurs
 // Transmutation des arguments
roid swap( int *x, int *y ) {
  int temp = *x;
  int temp =

*x = *y;

*y = temp;
int main() {
  int a = 10, b = 20;
  std::cout << setw(3);
  std::cout << a << b << std::endl;
swap( &a, &b );
std::cout << a << b << std::endl;
  return 0:
```

```
Paramètres références
 // Transmutation des arguments
roid swap( int &x, int &y ) {
  int temp = x;
  x = y;
y = temp;
int main() {
  int a = 10, b = 20;
  std::cout << setw( 3 );
   std::cout << a << b << std::endl;
swap( a, b );
std::cout << a << b << std::endl;
  return 0:
```

14728 ORD - Langage C++

Les deux exemples ci-contre sont deux versions d'un même programme. Le premier programme (celui de gauche) exploite l'appel par référence via pointeurs pour inverser le contenu des arguments a et b de main. La seconde version du programme (celui de droite) exploite l'appel par référence via références pour effectuer la même tâche.

Il est important de noter les distinctions dans ces deux versions du programme :

- 1. Les paramètres pointeurs requièrent l'opérateur de déréférencement (*) dans le corps de la fonction pour manipuler le contenu des arguments a et b, et l'invocation doit fournir les adresses de ses arguments à la fonction (d'où l'opérateur d'adresse & accompagnant les arguments).
- 2. Les paramètres références ne requièrent que l'utilisation de l'opérateur d'adresse (&) dans l'entête de la fonction afin de déclarer les paramètres comme des alias des arguments.



🄼 Arithmétique des pointeurs

- Une variable pointeur contient une adresse mémoire (d'une autre variable?)
 - On peut manipuler cette adresse comme si c'était une valeur entière

```
int t[] = { 10, 20, 30, 40 };
int *p = &t[ 0 ];
                                                        t : 10 20 30 40
for ( int i = 0; i < 4; i++ ) {
  cout << *p << " ";</pre>
```

 Puisque p est de type int *, p++ incrémente l'adresse dans p de la taille d'un int

Marco Lavoie

14728 ORD - Langage C++

Les pointeurs étant des adresses mémoire, ce sont des opérandes valides dans des expressions arithmétiques, d'affectation et de comparaison. Toutefois, pas tous ces opérateurs ne sont applicables aux pointeurs.

Les opérateurs arithmétiques applicables aux pointeurs sont limités: on peut les incrémenter (++) ou les décrémenter (--), leur additionner (+ ou +=) ou leur soustraire (- ou -=) un entier et leur soustraire un autre pointeur.

Supposons dans l'exemple ci-contre que le bloc mémoire associé au tableau t débute à l'adresse 3000. On aura donc sur une architecture 16 bits &t[0] = 3000, &t[1] = 3002, &t[2] = 3004, etc. Comme p est initialisé à 3000, l'expression p++ incrémente l'adresse dans p non pas de un octet, mais de 2 octets puisque p est déclaré comme pointeur de type int (un entier étant stocké dans 2 octets dans une architecture 16 bits). Donc la boucle incrémente successivement p à 3002, 3003, 3006, etc.



🄼 Arithmétique des pointeurs (suite)

Autre exemple :

```
const int N = 7;
int t[ N ] = { 10, 20, 30, 40, 50, 60, 70 };
for ( int *p = &t[ N-1 ]; p > &t[ 0 ]; p -= 2 )
  cout << *p << " ";</pre>
```

– Qu'affiche le code ci-dessus?

70 50 30

- p est décrémentée de 2 après chaque itération
- Aucune itération lorsque p == &t[0]

Marco Lavoie

14728 ORD - Langage C++

Dans l'exemple ci-contre, le pointeur p est initialisé à l'adresse du dernier élément du tableau t, et à chaque itération de la boucle for l'adresse dans p est décrémentée de 4 octets (puisqu'un entier est stocké sur deux octets et que p est un pointeur de type int, p -= 2 correspondant à une décrémentation de 4 octets de l'adresse contenue dans p).



🔼 Restrictions de type

· Le type d'une variable pointeur doit correspondre au type de la variable pointée

```
int *p1 - &v[ 0 ];  

double *p2 = &v[ 0 ];  

x
```

- Cette restriction est imposée par l'arithmétique des
 - Dans l'exemple ci -dessus, incrémenter p2 ne fait pas pointer au prochain élément de v, ce qui est le cas avec pl
- Cette restriction peut être contournée avec le transtypage explicite

```
double *p2 - reinterpret cast< double * >( &v[ 0 ] );
```

Auleur Marco Lavoie | Adaptation Sébastien Bois

25908 IFM Langage C++

Comme mentionné précédemment, une variable pointeur doit avoir un type qui correspond au type de valeurs pointées par celle-ci. Dans l'exemple ci-contre l'affectation de p2 à l'adresse du début du tableau v est invalide car le type de la variable pointeur (double) ne correspond pas au type de la valeur pointée (int).

Malgré le fait que le transtypage explicite permet de contourner cette restriction, il est très dangereux d'effectuer un tel transtypage dans un programme, au risque de causer une défaillance du programme à l'exécution. Généralement on évite de faire de tels transtypages en programmation, sauf dans des circonstances bien particulières et hors sujet de ce cours.



Relation entre pointeurs et tableaux

- Les tableaux et pointeurs sont intimement
 - Leur utilisation est presque interchangeable
- En effet, un tableau est considéré par le compilateur comme une variable pointeur
 - int t[] devient const int *t
 - L'opérateur [] est convertit en arithmétique des pointeurs équivalente

Marco Lavoie

14728 ORD - Langage C++

En C++, l'utilisation des tableaux et des pointeurs est presque interchangeable. On peut se représenter le nom d'un tableau comme étant un pointeur constant initialisé à l'adresse du premier élément du tableau. En fait, on peut remplacer l'opérateur d'indexation [] dans une par expression l'arithmétique de correspondant appliqué au nom du tableau, qui ajit comme pointeur.



🄼 Pointeurs vs tableaux (suite)

- Exemples
 - Supposons les déclarations suivantes

int v[10], u;

Les expressions suivantes sont équivalentes

p = &v[0];p = v;p = &v[3];p = v + 3;u = v[2];u = *(v + 2);p = v;p[5] = 17;void f(int []); void f(int *);

Marco Lavole

14728 ORD - Langage C++

Comme le démontre les exemples ci-contre, on peut traiter un tableau comme un pointeur et l'utiliser dans l'arithmétique des pointeurs. Par exemple, l'expression *(v + 2) renvoie à l'élément de tableau v[2]. En général, toute expression d'un tableau comportant des indices peut s'écrire avec un pointeur et un décalage du contenu de ce dernier.

Réciproquement, les pointeurs peuvent être indexés de la même façon qu'un tableau. L'avant dernier exemple ci-contre fait référence au sixième élément de v via p[5], et ce même si p est non pas un tableau mais un pointeur contenant l'adresse du tableau v.

Comme le nom d'un tableau est essentiellement un pointeur constant, il est interdit de changer l'adresse correspondant au début du tableau. Ainsi l'expression v += 2 serait interdite dans l'exemple ci-contre car elle tente de modifier l'adresse contenue dans v.



Pointeurs vs tableaux (suite)

- Exemples (suite)
 - Les programmeurs d'expérience interchangent souvent tableaux et pointeurs

```
const int N = 4;
int v[N] = \{ 10, 20, 30, 40 \};
for ( int i = 0; i < N; i++ )
  v[i] = v[i] + 10;
                                              équivalents
for ( int *p = v; p < v + N; p++ )</pre>
  *p = *p + 10;
```

14728 ORD - Langage C++

Les deux boucles ci-contre sont équivalentes. Le seconde boucle exploite un pointeur initialisé à l'adresse du tableau v pour parcourir les éléments de ce dernier. L'arithmétique des pointeurs permet de modifier l'adresse contenue dans p en l'incrémentant successivement, à chaque itération, afin de faire passer l'adresse qu'elle contient d'un élément de v au suivant.

Notez qu'on aurait aussi pu formuler la deuxième boucle for comme suit :

```
for (int *p = v, int i = 0; i < N; i++)
 p[i] = p[i] + 10;
```

Comme mentionné précédemment, l'opérateur d'indexation ([]) et celui de déréférencement (*, avec l'arithmétique des pointeurs) sont interchangeables dans la manipulation de tableaux comme des pointeurs ou vice-versa.



Exercice #5.1

- Sans faire appel à l'opérateur [] ni à une variable d'itération de type int, complétez le programme suivant de sorte qu'il affiche le nombre de valeurs paires trouvées dans le tableau
- Soumettez votre projet selon les indications de l'instructeur

```
int main() {
  const int N = 20;
  int valeurs[ N ];
  int compte = 0;
     // Remplir de valeurs aléatoires
srand( time( 0 ) );
for ( int i = 0; i < N; i++ ) {
valeurs[ i ] = rand() % 100;
std::cout << valeurs[ i ] << " ";
       ,
std::cout << std::endl;
```

std::cout << "# de valeurs paires: '
<< compte << std::endl;

Marco Lavole

14728 ORD - Langage C++

Pour solutionner cet exercice, vous devez modifier la boucle for de telle sorte qu'elle manipule le tableau valeurs comme un pointeur, exploitant l'arithmétique des pointeurs et l'opérateur de déréférencement pour remplacer toutes références à l'opérateur d'indexation dans la boucle. L'abandon de l'opérateur d'indexation rendra conséquemment caduc l'emploi de la variable d'itération i.

Inspirez-vous de l'exemple mentionné à la page précédente pour convertir la boucle for.



Tableaux et chaînes de caractères

· Une chaîne de caractères est stockée dans un tableau

```
char nom[] = "Yvan Larue";
char *ville = "Verdun";
```

Équivalentes à

```
char *ville = { 'V', 'e', 'r', 'd', 'u', 'n', '\0' };
```

 Les flux d'entrée/sortie peuvent manipuler des chaînes de caractères

char nom[30]; cin >> nom; cout << nom;

Marco Lavole

14728 ORD - Langage C++

Comme mentionné au chapitre 4, en C++ une chaîne est un tableau de caractères se terminant par le caractère nul ('\0'). On y accède par le biais du nom du tableau qui peut être considéré comme un pointeur constant vers le premier caractère de la chaîne. En ce sens, une chaîne est identique à un tableau puisqu'un nom de tableau est également un pointeur constant vers son premier élément.

On peut affecter directement une chaîne en provenance du clavier à un tableau en utilisant l'extraction de flux avec cin. Il est important de préciser que le tableau doit être assez grand pour stocker la chaîne entrée et son caractère nul de terminaison. Dans l'exemple ci-contre, la taille maximale de chaîne que peut fournir l'utilisateur pour nom est donc 29 caractères.



🄼 Affichage de chaînes de caractères

 Le flux de sortie cout affiche le contenu d'un tableau de char jusqu'au premier caractère nul ('\0') rencontré

 Il faut donc s'assurer qu'un '\0' sera trouvé dans la chaîne à afficher

Marco Lavoi

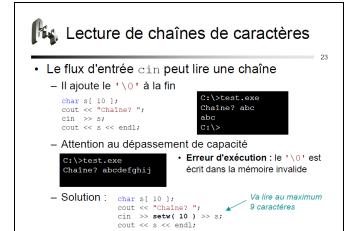
Marco Lavoie

14728 ORD - Langage C++

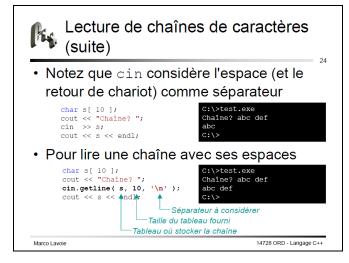
14728 ORD - Langage C++

Similairement au flux d'entrée cin qui peut lire une chaîne directement dans un tableau de caractères, le flux de sortie cout considère un tableau de caractères comme contenant une chaîne terminée par le caractère nul, et affiche donc les caractères contenus dans le tableau jusqu'au premier caractère nul rencontré dans celui-ci

Il est très important que tout tableau de caractères passé à cout contiennent au moins un caractère nul, sinon le flux de sortie affichera le contenu du tableau ainsi que le contenu de la mémoire vive suivant le bloc mémoire alloué au tableau, et ce jusqu'à ce qu'un caractère nul soit rencontré en mémoire. Il arrive même parfois qu'une telle situation cause l'arrêt de l'exécution du programme.



Comme mentionné précédemment, le tableau fourni à cin doit être assez grand pour stocker la chaîne entrée et son caractère nul de terminaison. On peut utiliser le manipulateur de flux setw pour assurer que la chaîne lue dans le tableau n'excède pas la taille de celui-ci. Notez que setw réserve une position du tableau pour le caractère nul de terminaison de chaîne.



Par défaut, cin considère le caractère blanc (i.e. l'espace) et le retour de chariot comme séparateurs de valeurs dans le flux d'entrée. Dans le premier exemple ci-contre, seuls les trois premiers caractères fournis par l'utilisateur (abc) sont lus dans le tableau s car le caractère blanc suivant c est considéré comme séparateur. Le reste des caractères fournis (def) demeurent dans le flux d'entrée pour la prochaine opération de lecture (c.à.d. pour le prochain cin exécuté).

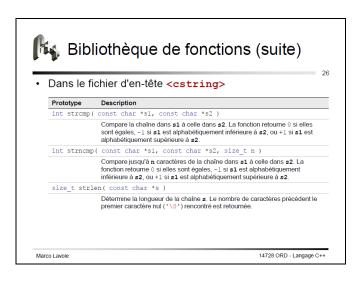
La fonction membre getline de cin permet de spécifier le séparateur à considérer lors de la prochaine lecture. Cette fonction est donc utile pour lire des chaînes contenant des blancs, tels qu'un nom complet ou une adresse postale.

Nous verrons au chapitre 6 ce qu'est une fonction membre.

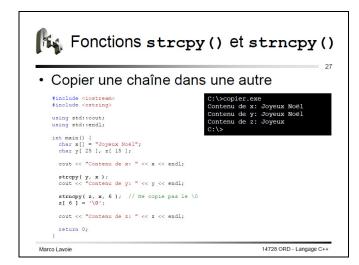


La bibliothèque de manipulation de chaînes offre de nombreuses fonctions utiles pour manipuler, comparer et rechercher des chaînes, les séparer en jetons (ou pièces logiques) et déterminer leur longueur.

Le tableau ci-contre n'énumère que les fonctions les plus couramment utilisées par les programmeurs. Plusieurs autres fonction non présentées ici sont disponibles.



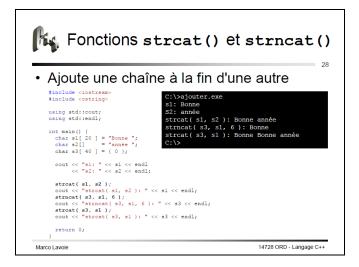
Plusieurs des fonctions de la bibliothèque, telles que certaines présentées ci-contre, contiennent des paramètres ayant le type de données size_t. Ce type est défini comme étant un type d'entier unsigned int ou unsigned long, dépendant du compilateur et du système d'exploitation. Les variables de type size_t peuvent donc être manipulées comme si elles étaient de type entier non signé.



La fonction **strcpy** copie son second argument, une chaîne, dans son premier argument, un tableau de caractères qui doit être assez grand pour stocker la chaîne et son caractère nul de terminaison.

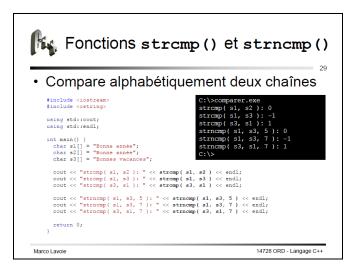
La fonction strncpy équivaut à strcpy, à la seule différence qu'elle spécifie le nombre de caractères maximal à copier de la chaîne au tableau. Cette fonction est pratique lorsqu'on ne connaît pas à priori la longueur de la chaîne qui sera fourni en second argument mais qu'on connaît la taille du tableau fourni en premier argument.

Sachez que la fonction **strncpy** ne copie pas nécessairement le caractère nul de terminaison de son deuxième argument, car ce caractère n'est copié que si le nombre de caractères à copier dépasse la longueur de la chaîne d'au moins un caractère.



La fonction streat ajoute son second argument, une chaîne, à la fin du contenu de son premier argument, un tableau de caractères contenant une chaîne. Le premier caractère du second paramètre remplace le caractère nul ('\0') qui termine la chaîne contenue dans le premier argument. Le programmeur doit s'assurer que le tableau servant à stocker la première chaîne est assez grand pour stocker la combinaison des deux chaînes et du caractère nul de terminaison (lequel est copié de la deuxième chaîne).

Quant à la fonction strncat, elle ajoute un nombre maximum donné de caractères de la deuxième à la première chaîne, puis un caractère nul de terminaison est ajouté au résultat.



La fonction stremp compare la première chaîne à la deuxième, caractère par caractère. Elle retourne 0 si les chaînes sont identiques en contenu et en longueur., une valeur négative si la première est inférieure à la deuxième, ou une valeur positive si la première chaîne est supérieure à la deuxième.

Pour essayer de bien comprendre ce que signifie qu'une chaîne est « supérieure » ou « inférieure » à une autre, il faut se référer à la table des codes ASCII. Comme nous l'avons vu au chapitre 2, chaque caractère en C++ est représenté par une valeur entière dictée par la table ASCII. La fonction stromp compare les deux chaînes caractère par caractère selon leurs codes ASCII respectifs. Dès que la fonction rencontre deux caractères ayant des codes ASCII différents, elle retourne une valeur numérique déterminée selon ces deux codes.



La fonction strlen prend une chaîne comme argument et renvoie le nombre de caractères inclus dans cette chaîne. Le caractère nul de terminaison n'est toutefois pas inclus dans la longueur.

En fait, la fonction strlen retourne l'indice de l'élément contenant le caractère nul dans la chaîne. Par exemples, dans un tableau t contenant une chaîne de trois caractères, le caractère nul sera contenu dans le $4^{\text{ième}}$ élément du tableau, soit dans t [3].

Exercice #5.2

- Sans faire appel aux fonctions du fichier d'entête <cstring>, définissez les deux fonctions suivantes
 - int maStrlen(const char *) retourne la longueur de la chaîne en paramètre
 - char * maStrReverse(char *) inverse le contenu de la chaîne en paramètre et retourne celleci comme valeur de retour
- N'oubliez pas les conventions d'écriture
- · Soumettez votre projet selon les indications de l'instructeur

14728 ORD - Langage C++

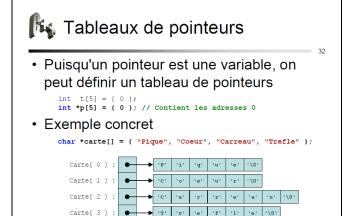
14728 ORD - Langage C++

Cet exercice vous demande de proposer du code source correspondant à la fonction strlen de la bibliothèque standard. Comme mentionné à la page précédente, la longueur d'une chaîne dans un tableau de caractères correspond à l'indice de l'élément du tableau contenant le caractère nul de terminaison de chaîne.

En ce qui concerne l'implantation de la fonction maStrReverse, vous pouvez exploiter une boucle for avec deux variables d'itération : la variable 10 est initialisée à 0 et incrémentée à chaque itération, alors que la variable hi est initialisée à la position précédent le caractère nul dans la chaîne et décrémentée à chaque itération. Tant que les deux indices (10 et hi) ne se sont pas rejoint, le code dans la boucle doit échanger les éléments s[lo] et s[hi].

L'entête d'une telle boucle peut être comme suit :

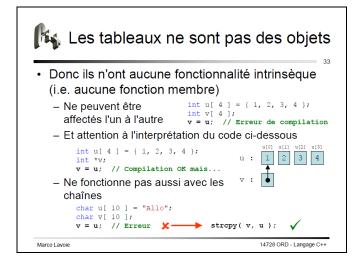
for (int lo=0, hi=strlen(s)-1; lo < hi; lo++, hi--)



Les tableaux peuvent également contenir des pointeurs. Une utilisation courante d'une telle structure de données consiste à former un tableau de chaînes, communément appelé tableau de chaînes de caractères. Chaque élément dans le tableau représente une chaîne de caractères, bien qu'en C++ une chaîne soit en fait un pointeur constant vers le premier caractère de celle-ci.

Dans l'exemple ci-contre, même si les chaînes semblent stockées dans le tableau carte, en fait seuls les pointeurs y sont stockés, chaque pointeur ciblant le premier caractère de sa chaîne correspondante (c.à.d. l'élément du tableau contient l'adresse de ce caractère).

Même si le tableau de pointeur est de taille fixe (c.à.d. qu'il ne peut contenir qu'un nombre limité de pointeurs), chacun de ses pointeurs peut pointer vers une chaîne de n'importe quelle longueur. Cette souplesse est un exemple des puissantes possibilités de structuration de données du C++.



Contrairement à d'autres langages comme Java et C#, les tableaux en C++ ne sont pas des objets (tels que les collections en Java). Conséquemment, plusieurs opérateurs communément exploités pour manipuler des tableaux dans ces langages ne fonctionnent pas de la même façon en C++. Il est par exemple impossible de copier une chaîne de caractères d'un tableau à un autre simplement en affectant le premier au second. Il faut obligatoirement exploiter la fonction strcpy ou strncpy pour réaliser un tel transfert.

Nous verrons au chapitre 16 la classe string qui offre des fonctionnalités de manipulation de chaînes comparables à celles retrouvées dans Java.



🔼 Erreurs de programmation

 Déréférencer un pointeur non initialisé ou initialisé à 0 cause une erreur d'exécution

```
- Exemple: int *v;
    cout << *v; // v pointe vers quoi?</pre>
```

 Initialiser un pointeur à une constante plutôt qu'à une adresse

```
- Exemple : \inf_{int \ v = 8; \atop int \ *p1, \ *p2;}
                                p1 = v;
p2 = &v;
```

 Confondre le contenu du pointeur et son contenu déréférencé (i.e. oublier l'opérateur *)

```
cout << *p2 << p2; // affiche 8 et l'adresse de v
```



Erreurs de programmation (suite)

- Exploiter l'arithmétique des pointeurs sur un pointeur ne correspondant pas à un tableau est généralement une erreur de logique
- Déborder la capacité d'un tableau via l'arithmétique des pointeurs
- Tenter de changer l'adresse contenue dans un tableau

```
- Exemple : \inf_{\substack{\text{int t[ 10 ], u[ 10 ];} \\ \text{int }^*v;}}
                           v = t; // OK

u = t; // Erreur: int u[ 10 ] équivalent à const int *u
```

Marco Lavoie

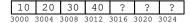
14728 ORD - Langage C++

Les pointeurs sont sources de multiples erreurs de programmation dues à la complexité de leur utilisation en C++. La cause d'erreur la plus fréquente est d'appliquer l'opérateur de déréférencement (*) à une variable pointeur contenant une adresse invalide, soit l'adresse nulle (c.à.d. 0), soit une adresse n'étant associé à aucun bloc mémoire alloué au programme. Une telle situation est une erreur de logique, c'est-à-dire que le programme est compilé avec succès, mais il ne fonctionnera pas correctement à l'exécution.

Une autre erreur logique fréquemment occasionnée par l'arithmétique de pointeurs contenant l'adresse d'un tableau est le débordement d'indice. La boucle suivante est un exemple d'une telle erreur :

```
int v[] = \{ 10, 20, 30, 40 \};
int *pv = &v;
for (int i = 0; i \le 4; i++)
  cout << *( pv + i ) << endl;
```

Lors de l'itération pour i = 4, *(pv + i) donne l'adresse de l'entier succédant à 40 dans le tableau (supposant une architecture 32 bits et $\mathbf{ev} = 3000$):



Pour i = 4, (pv + i) = (3000 + 4 * 4) = 3016, donc * (pv + i) retourne la valeur entière stockée au bloc de 4 octets débutant à l'adresse 3016 (c.à.d. ?).



🄼 Erreurs de programmation (suite)

- Ne pas allouer assez d'espace pour le caractère nul ('\0') dans un tableau de caractères
 - Et utiliser une chaîne dans un tableau de caractères n'étant pas terminée par '\0'
- Exploiter les fonction strncpy() et strncat() sans s'assurer que la chaîne résultante sera terminée par '\0'
- · Lire une chaîne avec cin en oubliant que l'espace est considéré comme un séparateur
 - Utiliser plutôt cin.getline() lorsque requis

Marco Lavoie

14728 ORD - Langage C++

La manipulation de chaînes dans des tableaux de caractères occasionne souvent des erreurs de logique. C'est le cas, entre autres, si la chaîne n'est pas terminée par le caractère nul dans le tableau, ce qui peut être le résultat d'un appel mal planifié à strncpy.

Toute manipulation de chaînes de caractères dans un programme doit susciter chez le programmeur une attention particulière afin d'éviter de telles erreurs de logiques, qui ne sont pas détectées à la compilation mais requièrent l'utilisation du débogueur pour être identifiées lors de l'exécution du programme.



🌃 Bonnes pratiques de programmation

Privilégier les paramètres de référence plutôt que les paramètres pointeurs

```
void swap( int &x, int &y ) {
  int temp = x;
 x = y;
y = temp;
```

- Toujours s'assurer qu'un tableau de caractères est assez grand pour contenir la chaîne et le caractère nul ('\0')
- Éviter l'arithmétique des pointeurs autant que possible
 - Privilégier l'opérateur [] à l'opérateur *

14728 ORD - Langage C++

Comme mentionné précédemment, la manipulation de pointeurs dans un programme C++ est souvent la cause d'erreurs de logique qui occasionnent la défaillance du programme lors de son exécution. Conséquemment le programmeur doit porter une attention particulière aux instructions d'un programme manipulant des pointeurs. Lors bonnes pratiques de programmation dans de telles circonstances évitent bien des heures de débogages au programmeur averti.



Devoir #4

- Solutionnez le problème distribué par l'instructeur
 - Créer un nouveau projet console Visual Studio
 - Solutionner le problème tel que décrit, avec les spécifications supplémentaires suivantes
 - · Le programme doit aussi transmettre de façon sonore le message
 - Dans la console, afficher une barre oblique (/) entre chaque mot
 - Les caractères autres que lettres et chiffres doivent être affichés dans la console sous forme d'espace et de façon sonore par une pause équivalente à celle entre deux mots
 - Le programme doit afficher le nombre de caractères invalides (i.e. n'ayant pu être transmis)
 - Respectez l'échéance imposée par l'instructeur
 - Soumettez votre projet selon les indications de l'instructeur
 - Attention : respectez <u>à la lettre</u> les instructions de l'instructeur sur la façon de soumettre vos travaux, s*inon la note EC sera attribuée* à

Marco Lavole

14728 ORD - Langage C++

Votre programme doit lire des caractères du clavier (à l'aide de cin) et afficher ceux-ci en caractères Morse correspondants. Vous pouvez utiliser à cette fin une structure de sélection switch imbriquée dans une boucle parcourant le tableau de caractères lus à l'aide de l'instruction cin.getline.



Devoir #4 (suite)

 Voici quelques notes supplémentaires sur le devoir

- L'instructuer a distribué un exécutable de la solution. Inspirez-vous en!
- Le fichier d'en-tête <windows.h> contient les deux fonctions suivantes qui vous seront utiles
 - Beep(int frequence, int duree)
 - Active le haut-parleur de l'ordinateur à une fréquence donnée pour une durée donnée en millisecondes
 - Exemple: Beep (500, 50); // ton de 500 Hz pour 50 ms
 - Sleep(int duree)
 - Suspend temporairement l'exécution pour une durée donnée en millisecondes
 - Exemple: Sleep(100); // Arrêt d'exécution 100 ms

Marco Lavoie

14728 ORD - Langage C++

Alternativement au switch, vous pouvez définir un tableau de 39 rangées (une pour chaque caractère traduisible) et de 6 colonnes (maximum de ti et ta dans un caractère traduisible). Ce tableau étant initialisé à la déclaration à l'aides des constantes TI et TA, vous pouvez y faire référence dans le code afin de convertir les caractères en code Morse sonore :

```
const int TI = 50; // durée d'un ti en ms const int TA = 3 * TI; // durée d'un ta en ms
                              // durée d'un ti en ms
// ti et ta pour chaque caractère valide
const int codeMorse[][ 6 ] = {
   { TI, TA },
  { TA, TI, TI, TI }, // B
{ TA, TI, TA, TI }, // C
  { TA, TI, TI },
                             // D
   { TI },
```



Devoir #4 (suite)

- Notes supplémentaires (suite)
 - Pour avoir plus de détails sur la transmission en code Morse, consultez Wikipédia

(http://fr.wikipedia.org/wiki/Alphabet morse)

· Surtout la section intitulée « Représentation et cadence »

14728 ORD - Langage C++

N'hésitez pas à consulter la référence ci-contre pour bien comprendre les principes de base du code Morse.

La solution de ce devoir exigera de votre part plusieurs heures de travail et de débogage. Il est donc important que vous commenciez à y travailler le plus tôt possible afin d'avoir suffisamment le temps de demander de l'aide à l'instructeur au besoin.



🄼 Pour la semaine prochaine

- · Vous devez relire le contenu de la présentation du chapitre 5
 - Il y aura un examen sur les chapitres 1 à 5 au prochain cours
 - · À livres et ordinateurs ouverts
 - · Accès au compilateur interdit
 - Profitez-en pour réviser le contenu de ces chapitres

Marco Lavole

14728 ORD - Langage C++

En début de classe la semaine prochaine vous aurez à répondre à des questions sur le C++ sans consultation du matériel pédagogique. Vous êtes donc fortement encouragé à relire les notes de cours du chapitre 5.

Profitez-en pour réviser le contenu des chapitres précédents.