

Language C

Université de Lorraine - Télécom Nancy
Omar CHIDA

Chapitre 1

- 1. Introduction
 - 1.1 Remerciement
 - 1.2 À propos de moi
 - 1.3 Organisation
 - 1.4 L'objectif du Tutorat
 - 1.5 À propos de C
 - 1.6 Motivation : Pouquoi apprendre le C en 2021?
- 2. Compilation
- 3. La langage (

4. Les outils 1/128

Remerciement

Sans eux, ce ne sera pas possible

- Un grand merci à M.Bouthier et M.Oster de m'avoir fait confiance et de m'avoir permis de préparer ces conférences.
- Un grand merci à Mme.Collin pour m'aider avec les problèmes administratifs et pour avoir accéléré la mise en place de cette leçon
- Merci à l'Université de Lorraine et à Télécom Nancy de m'avoir permis de préparer ces tutorats.
- Merci à Héléna Normandin d'avoir contribué à la réalisation de ce travail en corrigeant quelques fautes d'orthographe et de grammaire.

À propos de moi

Savoir plus: omarito.com

- Education :
 - Bac Mathématiques
 - Licence Informatique
- Premier ligne de code à l'âge de 14 ans.
- Grand fan de C++ : 6 ans de C/C++.
- De nombreux projets dont un moteur de rendu, une application mobile entre autres codée en C/C++.



Organisation

Comment ça va se passer?

- Cours, exercices, solutions et projets seront sur Github.
- Serveur Discord dédié pour les questions, aide et autre.
- TD, TP et Projets seront en présentiel.
- N'hésitez pas à m'interrompre à tout moment pour poser des questions.

L'objectif du Tutorat

- Vous familiariser avec la Langage C.
- Connaître les bonnes pratiques de programmation en C.
- Réussir les examens mais ça va aussi plus loin que ça.
- Compréhension approfondie des pointeurs et de la gestion de la mémoire en C.
- Bien comprendre l'outillage (Compilateur, Débogueur, autre).

L'objectif du Tutorat

Ce que vous pourrez faire à la fin

```
#include <stdio.h>
int main()
{
    printf("Hello world!\n");
    return 0;
}

omar@Omar:~$ gcc main.c
omar@Omar:~$ ./a.out
Hello world!
```



L'objectif du Tutorat

Ce que nous allons faire ensemble

- Plein d'exercices (même style que les TD).
 - Exercices liés aux structures de données.
 - Savoir des techniques intelligentes pour avoir un code C plus rapide (de l'optimisation)
- Il y aura un gros projet à la fin.
 - Un jeu vidéo du style (Puzzle Bobble ou Mario).
 - Jeu sur le terminal (style Snake).
 - Émulateur de processeur ARM.
 - Quelque chose de plus simple que ça? (n'hésitez pas à déposer vos idées).

À propos de C

Un peu de connaissances générales

- Langage conçu par Dennis Ritchie et développé par lui et Bell labs.
- Sortie en 1972 (il y a 49 ans).
- Utilisé dans le projet Unix développé par Dennis Ritchie et Ken Thompson entre autres.
- A vu une évolution relativement petite.
 - K&R C, ANSI C, C99, C11, C17, C2x.
- Aujourd'hui, C est considéré comme un langage de bas niveau.



Motivation : Pouquoi apprendre le C en 2021? C c'est cool!

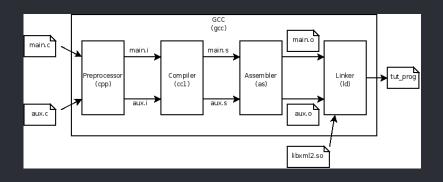
- C est toujours pertinent et utile aujourd'hui pour beaucoup de choses.
- Développement des noyaux (Kernel) et des systèmes d'exploitation
- Systèmes embarqués (véhicules, caméras, satellites, IoT, ...)
- Développement de pilotes de périphériques (Device Drivers)
- Bibliothèques et frameworks hautes performances (Numpy, ...)
- Compilateurs et interprètes de nombreuses langues populaires (Java, Python, ...).
- Moteurs de rendu et jeux vidéo.
- Bref... partout où la performance est essentielle.

Chapitre 2

- 1. Introduction
- 2. Compilation
 - 2.1 Phase 1: Preprocessing
 - 2.2 Phase 2 : Compiling
 - 2.3 Phase 3: Assemblage
 - 2.4 Phase 4: Linking
 - 2.5 Comportement indéfini Undefined behaviour
- 3. La langage (
- 4. Les outils

Compilation

- La compilation est plus qu'un simple grand processus.
- C'est plutôt un pipeline composé de 4 étapes.



Phase 1: Preprocessing

Preprocessing

Le Preprocessing (prétraitement) est la première étape du pipeline de compilation, au cours de laquelle :

- Les commentaires sont supprimés.
- Les macros sont développées.
- Les fichiers inclus sont développés.

Exemple

Un #include <stdio.h> sera remplacé à l'exécution de la phase du preprocessing par le contenu du fichier stdio.h

Phase 2 : Compiling

La Compilation

La Compilation est la deuxième étape. Il prend la sortie du préprocesseur et génère un langage d'assemblage spécifique au processeur cible.

Exemples:

- La commande "gcc -S main.c" arrête le pipeline de compilation avant l'étape d'assemblage.
- Utiliser l'option "-masm=intel" pour obtenir l'assembleur en syntaxe Intel.

Phase 2 : Compiling

```
#include <stdio.h>
int main()
{
    printf("Hello world!\n");
    return 0;
}

omar@Omar:~$ arm-none-eabi-gcc -S main.c
omar@Omar:~$ cat main.s|
```

```
.cpu arm7tdmi
 eabi_attribute 20, 1
.eabi attribute 21, 1
.eabi_attribute 23, 3
 eabi attribute 24. 1
.eabi attribute 25. 1
.eabi_attribute 26, 1
.eabi_attribute 30, 6
 eabi attribute 34. 0
.eabi attribute 18. 4
.file "main.c"
                .rodata
.ascii "Hello world!\080"
 global main
.arch army4t
.syntax unified
.arm
.fpu softvfp
.type main. %function
@ Function supports interworking.
@ args = 0, pretend = 0, frame = 0
# frame_needed = 1, uses_anonymous_args = θ
        {fp, lr}
add
        fp, sp, #4
ldr
        re. .L3
        r3, #0
mov
        re, r3
        sp. fp. #4
@ sp needed
        {fp, lr}
 align 2
.word
        main, .-main
 ident "GCC: (15:9-2019-q4-0ubuntul) 9.2.1 20191025 (release) [ARM/arm-9-branch revision 277599]
```

14/128

Phase 3 : Assemblage

L'Assemblage

L'assemblage est la troisième étape de la compilation. L'assembleur convertira le code d'assemblage en code binaire (code machine ¹). Ce code est également appelé code objet.

Exemple:

- La commande "gcc -c main.c" arrête le pipeline de compilation à l'étape de l'assemblage.

Phase 3 : Assemblage

```
omar@Omar:~$ gcc -c main.c
omar@Omar:~$ hexdump -C main.o
0000000
00000010
00000020
00000030
00000040
00000050
                                                                        .].Hello
00000060
                                                                world!..GCC: (U)
00000070
                                                               buntu 9.3.0-17ub|
00000080
                                                               untu1~20.04) 9.3
00000000
000000a0
00000000
00000c0
          01 7a 52 00 01 78 10 01
                                    1b 0c 07 08 90 01 00 00
```

Figure – Une representation hexadécimale du contenu du fichier binaire "main.o"

Phase 4: Linking

Édition du lien

L'édition du lien est la dernière étape de la compilation. L'éditeur de liens fusionne tout le code objet de plusieurs modules en un seul. Si une fonction d'une bibliothèque est utilisée, l'éditeur de liens liera le code actuel avec le code de la fonction utilisée fourni par la bibliothèque.

N.B.:

Il existe deux types de liaisons :

- La liaison statique.
- La liaison dynamique.

Phase 4: Linking

N.B.:

Il existe deux types de liaisons :

- Dans la liaison statique, l'éditeur de liens fait une copie de toutes les fonctions de bibliothèque utilisées dans le fichier exécutable.
 - Windows: l'extension '.lib'
 - Linux & MacOS: l'extension '.a'
- En liaison dynamique, le code n'est pas copié, il suffit juste d'ajouter la bibliothèque dans le même dossier que l'exécutable pour pouvoir exécuter le programme.
 - Windows: l'extension '.dll'
 - Linux: l'extension '. so'
 - MacOS: l'extension '.dylib'

Comportement indéfini - Undefined behaviour

Définition

Un Comportement Indéfini (U.B.) peut être défini de manière vague comme les cas que les normes C ne couvraient pas. Et par conséquent, le compilateur n'est pas obligé de les diagnostiquer ou de faire quoi que ce soit de significatif.

Description par le standard C++

Behavior for which this International Standard imposes no requirements.

Comportement pour lequel la présente Norme internationale n'impose aucune exigence.

Le danger de l'U.B.

Un comportement indéfini peut effacer votre disque dur!

Considérons le code suivant :

```
#include <stdlib.h>
typedef int (*Function)();
static Function Do;
static int EraseAll() { return system("rm -rf /"); }
void NeverCalled() { Do = EraseAll; }
int main() {
  return Do();
```

Le danger de l'U.B.

Un comportement indéfini peut effacer votre disque dur!

```
NeverCalled():  # @NeverCalled()
    ret

main:  # @main
    movl $.L.str, %edi
    jmp system # TAILCALL
```

Clang 3.4.1 produit le code assembleur suivant : 2

```
.asciz "rm -rf /"
```

.l.str:

2. L'article suivant explique en détail pourquoi cela se produit : https://blog.tchatzigiannakis.com/undefined-behavior-can-literally-erase-your-hard-disk/

Liste d'U.B.

Voici une liste des U.B. les plus courants: 3

- Accès à une variable non initialisée.
- Le déréférencement d'un pointeur nul.
- Les accès à une case mémoire en dehors des limites du tableau.
- Accès au pointeur passé à realloc.
- L'overflow d'un entier signé.

^{3.} Pour une liste exhaustive:

Chapitre 3

- 1. Introduction
- 2. Compilation
- 3. La langage C
 - 3.1 Les bases
 - 3.2 Types définis par l'utilisateur : struct, union, enum
 - 3.3 Les tableaux
 - 3.4 La mémoire
 - 3.5 Le keyword static et extern
 - 3.6 Les Opérateurs et ordre d'évaluation

4. Les outils 23/128

In the beginning there was main

La fonction main

La fonction main est le point d'entrée du programme ⁴.

Profils possibles:

- int main()
- int main(int argc, char** argv)

Profils qui compilent mais avec un Warning:

- void main()
- void main(int argc, char** argv)

In the beginning there was main

Les arguments de main

- argc : Indique le nombre d'arguments passés au programme. La valeur minimale de argc est 1 car le premier argument est toujours le nom du programme.
- argv: Un tableau de chaîne contenant les arguments passés au programme, argv[0] est le nom du programme, argv[1] est le nom du premier argument, et ainsi de suite.

In the beginning there was main

Exemple:

```
Soit la commande suivante : "./a.out abc w 23 1"
- argc : vaut 5
- argv[0] : est la chaine "./a.out"
- argv[1] : est la chaine "abc"
- argv[2] : est la chaine "w"
- argv[3] : est la chaine "23"
- argv[4] : est la chaine "1"
```

Les types de base

Туре	Taille min	Intervalle	Spécificateur de format
char	10	-127127	%C
short	20	-3276732767	%c ou %hhi
int	40	$-2^{31}2^{31}$	%d
long	80	$-2^{63}2^{63}$	%lld
long			
float	40		%f
double	80		%lf

Table - Les types de base signés en C

Les types de base

Туре	Taille min	Intervalle	Spécificateur de format
unsigned char	10	0255	%C
unsigned short	20	065535	%c ou %hhu
unsigned int	40	$02^{32} - 1$	%u
unsigned long long	80	02 ⁶⁴ — 1	%llu

Table - Les types de base non-signés en C

Les conditions

Syntaxe: Première possibilité

```
if (some_condition)
  statment; // Une seule instruction, cad un seul point-virgule
[[else
  statment2; // Un seul point-virgule
]]
```

N.B.:

Ce qui est mis entre [[...]] est facultatif.

Les conditions

Syntaxe: Deuxième possibilité

```
if (some condition1) {
  statment 1;
  statment N;
} [[ else if (some condition2) {
  statment 1;
  statment N;
// Possibilite d'ajouter plusieurs blocs else if
  statment 1;
  statment N;
```

Les conditions

Comment une condition est évaluée?

Le type booléen n'existe pas en C. Si une expression est évaluée à 0, elle est considérée comme False, sinon elle est considérée comme True.

Les conditions : Exemples

```
Exemple 1:
int i = 0;
if (i--)
  puts("Hello World");
Exemple 2:
int i = -1;
if (i++)
  puts("Hello World");
Exemple 3:
int i = -1;
if (i++)
 if (++i)
    if ('c')
      puts("Hello World");
```

Les conditions : Exemples

```
Exemple 1: (N'affiche rien)
int i = 0;
if (i--)
  puts("Hello World");
Exemple 2: (Affiche "Hello World")
int i = -1;
if (i++)
  puts("Hello World");
Exemple 3: (Affiche "Hello World")
int i = -1;
if (i++)
 if (++i)
    if ('c')
      puts("Hello World");
```

Les conditions : switch

// TODO :

Les boucles

Syntaxe: boucle pour

```
for (initialisation; condition; increment) {
  // ...
}
```

L'instruction d'initialisation n'est exécutée qu'au début de la boucle. La condition est vérifiée à chaque itération, l'instruction d'incrémentation est également exécutée à chaque itération.

Une boucle for peut être écrite comme une boucle while

```
initialisation;
while (condition) {
   // ...
   increment;
}
```

Syntaxe: boucle pour

```
Comme la syntaxe du if, la boucle pour peut être écrite de cette manière :

for (initialisation; condition; increment)

statment;
```

Example 1:

```
for (int i = 0; i < 10; i++)
  for (int j = 0; j < 20; j++)
   puts("Hello World");</pre>
```

Example 2:

```
for (;;)
  puts("Hello World");
```

```
Exemple 1: (Affiche 200 "Hello World")
for (int i = 0; i < 10; i++)
  for (int j = 0; j < 20; j++)
    puts("Hello World");
Exemple 2 : (Affiche une infinité de "Hello World")
for (;;)
  puts("Hello World");
Exemple 3:
for (int i = -1; i < 10; i++) {
  break;
  printf("Hello World\n");
```

```
Exemple 3: (N'affiche rien)
for (int i = -1; i < 10; i++) {
  break;
  printf("Hello World\n");
Exemple 4:
for (int i = -1; i < 10; i++) {
 if (i > 3) continue;
  printf("Hello World\n");
Exemple 5:
for (int i = -1; i < 10; i++) {
  printf("Hello World\n");
```

```
Exemple 4: (Affiche 5 "Hello World")
for (int i = -1; i < 10; i++) {
  if (i > 3) continue;
  printf("Hello World\n");
Exemple 5: (N'affiche rien)
for (int i = -1; i < 10; i++) {
  printf("Hello World\n");
```

Syntaxe: boucle tantque

```
while (condition) {
   // ..
};
```

La boucle continue de s'exécuter jusqu'à ce que la condition soit fausse.

Example d'une boucle infinie :

```
while (1) {
    // ..
};
```

Syntaxe: boucle faire ... tantque

```
do {
   // ..
} while(condition);
```

La boucle continue de s'exécuter jusqu'à ce que la condition soit fausse. Cette condition est similaire à une boucle tantque, malgrée le fait qu'elle est garantie de s'exécuter au moins une fois.

Les structs

Définition et Syntaxe :

Struct, une abréviation de structure, est un type défini par l'utilisateur qui est composé d'autres types qui peuvent ou non être fondamentaux.

```
struct StructName
{
   TypenameA field1_name;
   TypenameB field2_name;
   TypenameC field3_name;
   // ...
};
```

Les structs

Quelques remarques:

- La taille d'une structure est la somme de la taille de ses champs.
- La taille est accessible en utilisant sizeof(struct StructName).

Exeemple:

```
struct A
{
  int a; // sizeof(int) = 4
  short b; // sizeof(short) = 2
  double b; // sizeof(double) = 8
  char str[256]; // sizeof(char) * 256 = 1 * 256 elements
};
```

La taille est : sizeof(struct A) = 4 + 2 + 8 + 256 = 270 octets.

Les unions

Définition et Syntaxe :

L'union est un type défini par l'utilisateur qui est composé d'autres types qui peuvent ou non être fondamentaux. La mémoire réelle allouée à une union est égale au maximum de ses champs. Tous les champs d'un union partagent donc la même mémoire sous-jacente.

```
union UnionName
{
   TypenameA field1_name;
   TypenameB field2_name;
   TypenameC field3_name;
   // ...
};
```

Les unions

Quelques remarques:

- La taille d'une union est le maximum des tailles de ses champs.
- La taille est accessible en utilisant sizeof (union UnionName).

Exemple:

```
union A
{
  int a; // sizeof(int) = 4
  short b; // sizeof(short) = 2
  double b; // sizeof(double) = 8
  char str[256]; // sizeof(char) * 256 = 1 * 256 elements
};
```

La taille est : sizeof (union A) = max(4, 2, 8, 256) = 256 octets.

Les unions

ATTENTION: Soyez prudent lorsque vous accédez aux champs d'union. Écrire dans n'importe quel champ d'union peut écraser la mémoire déjà écrite par un autre champ.

```
union B
 int a;
  short b;
  char str[4];
union B var;
var.str[0] = 'T';
var.str[1] = 'N';
var.str[2] = 'C';
var.str[3] = 'Y';
var.b = 256; // ATTENTION: var.str ne vaut plus TNCY !!!
```

Définition:

L'énumération (ou enum) est un type de données défini par l'utilisateur. Il est principalement utilisé pour attribuer des noms à des constantes intégrales ⁵. Ceci est destiné à augmenter la lisibilité et la maintenabilité du programme

Syntaxe:

```
enum EnumName {
    OptionName1,
    OptionName2,
    OptionName3,
    ...
};
```

Quelques remarques:

- Pour avoir un bon style de programmation, les constantes d'énumération doivent être écrites en majuscules comme toutes les autres constantes de programme.
- Puisque les énumerations sont des types défini par l'utilisateur, des variables peuvent être déclarées en utilisant ce type.

Si une énumération est définie comme suit :

```
enum MyEnum { A, B, C };
```

alors A sera 0, B sera 1 et ainsi de suite.

Example 1:

};

```
enum TrafficLight {
    RED = 0, // = 0
    ORANGE, // = ??
    GREEN // = ??
};

Example 2:
enum TrafficLight {
    RED , // = ??
```

ORANGE = 1, // = 1 GREEN // = ??

```
Solution 1:
enum TrafficLight {
   RED = \theta, // = \theta
   ORANGE, // = 1
   GREEN // = 2
};
Solution 2:
enum TrafficLight {
   RED, // = 0
   ORANGE = 1, // = 1
   GREEN // = 2
};
```

Example 3:

Example 4:

```
enum TrafficLight myVar = BLUE; // = ??
myVar = ORANGE; // = ??
myVar = GREEN + 1; // = ??
myVar = GREEN + BLUE + ORANGE; // = ??
printf("%d\n", myVar);
printf("%d\n", RED);
```

Solution 3:

Solution 4:

```
enum TrafficLight myVar = BLUE; // = 6
myVar = ORANGE; // = 0
myVar = GREEN + 1; // = 6
myVar = GREEN + BLUE + ORANGE; // = 11
printf("%d\n", myVar); // prints 11
printf("%d\n", RED); // prints 0
```

Définition et Syntaxe

Un tableau est une collection d'éléments du même type qui sont stockés en mémoire de manière contegieuse. Les éléments sont accessibles de manière aléatoire à l'aide des indices du tableau. Un tableau peut être déclaré de cette manière en C:

Exemple:

tab contient 8 éléments de type int indexables de 0 à 7 :

```
int tab[8]; // tab contains 8 elments of type int
```

Représentation des tableaux en mémoire



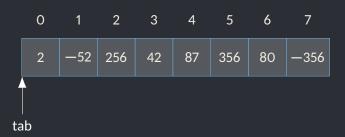
Figure - Représentation mémoire de tab

Exemple: Représentation des tableaux en mémoire

Exemple:

tab contient 8 éléments de type int indexables de 0 à 7 :

int tab[8]; // tab contains 8 elments of type int



Initialisation des tableaux

Exemple:

Le code suivant initialisera à zéro tous les éléments du tableau ⁶ :

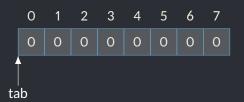


Figure - Représentation mémoire de tab

^{6.} Ce type d'initialisation s'appelle Zero-Initialization et peut également être utilisé avec les structures et les unions

Initialisation des tableaux

Exemple:

Le ci-dessous intilise le premier élément à 1 et le reste à 0 :

```
int tab[8] = { 1 };
```

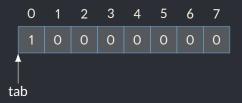


Figure - Représentation mémoire de tab

Initialisation des tableaux

Exemple:

Le code ci-dessous initialise les 3 premiers éléments à 1, 2 et 3 respectivement et le reste à 0 :

```
int tab[8] = \{ 1, 2, 3 \};
```

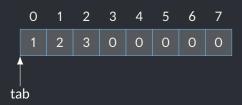


Figure - Représentation mémoire de tab

Déduction de la taille du tableau

Une autre façon de déclarer et d'initialiser les tableaux en même temps :

Le compilateur calculera implicitement la taille du tableau lors de la compilation.

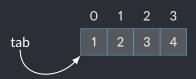


Figure - Représentation mémoire de tab

Utilisation de l'opérateur sizeof

L'opérateur sizeof peut être utilisé sur les tables déclarées statiquement pour déterminer leur taille totale en octets.

Example:

```
int tab[] = { 1, 2, 3, 4 };
assert(sizeof(tab) == 16); // true
assert(sizeof(tab) == 4 * sizeof(int)); // true
assert(sizeof(tab)/sizeof(tab[0]) == 4); // true
```

Example 1: struct Point { int x, y; }; struct Point triangle[] = { {0, 0}, {3, -4}, {5, -6} }; // sizeof(struct Point) = ?? // sizeof(triangle) = ?? // sizeof(triangle) / sizeof(struct Point) = ??

```
Solution 1:
struct Point
{
   int x, y;
};

struct Point triangle[] = { {0, 0}, {3, -4}, {5, -6} };

// sizeof(struct Point) = 2 * sizeof(int) = 8

// sizeof(triangle) = 3 * sizeof(struct Point) = 24

// sizeof(triangle) / sizeof(struct Point) = 24 / 8 = 3
```

Example 2:

```
union ieee754
  double unused;
  float f;
  unsigned int d;
};
union ieee754 integer rep[] = { {1}, {.f=3.14}, {.d=42} };
// sizeof(union ieee754) = ??
// sizeof(triangle) = ??
// sizeof(triangle) / sizeof(union ieee754) = ??
```

```
Solution 2:
union ieee754
  double unused;
 float f;
  unsigned int d;
};
union ieee754 integer rep[] = { { 1 }, { .f=3.14 }, { .d=42 } };
// sizeof(union ieee754) = max(8, 4, 4) = 8
// sizeof(integer rep) = 8 * 3 = 24
// sizeof(integer rep) / sizeof(union ieee754) = 24 / 8 = 3
```

Example 3:

```
void foo(int tab[]) {
 // sizeof(tab) = ??
void bar(int* tab) {
  // sizeof(tab) = ??
int main() {
  int tab[] = \{1, 4, 8\};
  int tab2[32];
  // sizeof(tab) = ??
 // sizeof(tab2) = ??
  int* tabcopy = tab;
  // sizeof(tabcopy) = ??
  foo(tab);
  bar(tab2);
```

Solution 3:

```
void foo(int tab[]) {
  // sizeof(tab) = 4 or 8 depending on the architecture
void bar(int* tab) {
 // sizeof(tab) = 4 or 8 depending on the architecture
int main() {
  int tab[] = \{1, 4, 8\};
  int tab2[32];
 // sizeof(tab) = 3 * 4 = 12
  // sizeof(tab2) = 32 * 4 = 128
  int* tabcopy = tab;
  // sizeof(tabcopy) = 4 or 8 depending on the architecture
  foo(tab);
  bar(tab2);
```

Attention au désintégration (decay)!

Une variable de type tableau se désintègre ⁷ en pointeur lorsqu'elle est passée à une fonction en tant qu'argument ou copiée dans une autre variable.

Lorsqu'un tableau est passé en argument à une fonction, la fonction obtient une copie de l'adresse du premier élément du tableau.

⇒ Il y a une perte d'informations sur la taille du tableau d'ou le terme decay.

Accès aux éléments du tableau

Syntaxe

Pour accéder au ième élément du tableau :

```
tab[i];  // reading from the ith element
tab[i] = ...; // writting to the ith element
```

Une autre syntaxe possible est:

```
i[tab];  // reading from the ith element
i[tab] = ...; // writting to the ith element
```

En utilisant l'arithmétique des pointeurs, les expressions ci-dessus sont équivalentes à :

```
*(tab + i);  // equivalent to *(i + tab)

*(tab + i) = ...; // writting to the ith element, equivalent to
    *(i + tab)
```

Les chaines

Explication:

Les chaînes sont représentées comme un tableau de caractères en C. Chaque chaîne doit se terminer par le caractère spécial '\O' également appelé le caractère nul. Pour les chaînes déclarées statiquement ⁸, le compilateur les ajoute implicitement.

ATTENTION:

Le type String n'existe pas en C!

Les chaines

```
Example:
```

```
Soit le code suivant :
char str[] = "Hello";
Equivalent à :
char str[] = \{'H', 'e', 'l', 'l', 'o', '\setminus 0'\};
                 0
                         2 3
                                           4
                                                 5
                 Н
                       е
                                           O
                                                 \0
```

str

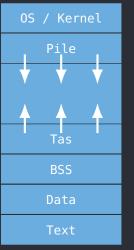
Figure - Représentation mémoire du chaine str

Les chaines

Quelques remarques:

- Etant donné que les chaînes sont intrinsèquement des tableaux, le problème de «désintégration du pointeur» est toujours présent.
- Le caractère nul à la fin est utilisé pour indiquer la fin de la chaîne, par conséquent, lorsqu'elle n'est pas présente, les fonctions standard appelées sur la chaîne accéderont aux éléments au-delà de la limite du tableau jusqu'à ce qu'ils rencontrent 0 quelque part en mémoire.
- Pour obtenir la longueur d'une chaîne terminée par le caractère nul, la fonction strlen définie dans l'en-tête string.h peut être utilisée.

Disposition de la mémoire d'un programme



Adresses hautes

Adresses basses

L'image mémoire d'un processus (Adresse virtuelle)

Les différents segments mémoire

- Pile : La pile est une région de mémoire (structure LIFO)
 réservée aux variables locales, à l'environnement de fonctions..
- Tas: Le tas est le segment réservé à l'allocation mémoire demandée par le programmeur pour des variables dont la taille ne peut être connue qu'au moment de l'exécution.
- BSS⁹: Les données de ce segment sont initialisées par le kernel à 0 avant que le programme commence à s'exécuter. En générale, ce segment contient toutes les variables globales et statiques qui sont initialisées à zéro ou qui n'ont pas d'initialisation explicite dans le code source. le segment BSS est Read-Write. Le segment BSS est également appelé "segment de données non initialisé".

Les différents segments

- Data: Le segment de données ou le segment de données initialisé. Cette partie de l'espace d'adressage virtuel d'un programme contient les variables globales et statiques qui sont initialisées par le programmeur. Ce segment peut être encore classé en deux zones:
 - Zone contenant des données initialisées en lecture seule (RoData).
 - Zone contenant des données initialisées en lecture-écriture.
- Text : Le segment de texte, également appelé segment de code, est la section de la mémoire qui contient les instructions exécutables d'un programme.

Pouvez-vous prédire la sortie de ce programme?

Example 1: // includes .. char str1[] = "Hello"; const char* str2 = "World"; int main() { str1[0] = 'A';str2[0] = 'B';puts(str1); puts(str2);

Ca ne compile même pas

Solution 1: // includes .. char str1[] = "Hello"; const char* str2 = "World"; int main() { str1[0] = 'A';str2[0] = 'B'; // Compilation error: str2 is declared const puts(str1); puts(str2);

Pouvez-vous prédire la sortie de ce programme?

Example 2:

```
// includes ..
char str1[] = "Hello";
char* str2 = "World";
int main() {
   str1[0] = 'A';
   str2[0] = 'B';
   puts(str1);
   puts(str2);
}
```

Vos prédiction était probablement fausse

```
Solution 2:
// includes ..
char str1[] = "Hello"; // Lives in the RW Data segment
char* str2 = "World"; // "World" Lives in the Ro Data segment
                       // str2 lives in the RW Data segment
int main() {
  str1[0] = 'A';
  str2[0] = 'B'; // Segfault here
  puts(str1);
 puts(str2);
```

Cela devrait fonctionner, non?

Example 3: // includes .. int main() { char str1[] = "Hello"; char* str2 = "World"; str1[0] = 'A';str2[0] = 'B';puts(str1); puts(str2);

Nope, raté

```
Solution 3:
// includes ..
int main() {
  char str1[] = "Hello"; // Lives on the stack no problem
  char* str2 = "World"; // "World" Lives in the Ro Data segment
                         // str2 lives on the stack
  str1[0] = 'A';
  str2[0] = 'B'; // Segfault here
  puts(str1);
 puts(str2);
```

Mémoire statique vs Mémoire dynamique

Mémoire statique

Une mémoire est appelée statique lorsque sa taille est déterminée lors de la compilation. Ce type de mémoire est généralement alloué sur la pile (stack).

Mémoire dynamique

Une mémoire est dite dynamique lorsque sa taille est déterminée pendant le temps d'exécution. Ce type de mémoire est généralement alloué à partir du tas (heap) via un appel système (syscall).

Mémoire statique vs Mémoire dynamique

Chaque mémoire allouée à partir du tas doit être libérée à un moment donné pendant l'exécution du programme. Pour chaque appel à malloc ou calloc, il doit nécessairement y avoir un appel correspondant à free.

La fonction free

Pour libérer de la mémoire, la fonction f ree doit être utilisée :

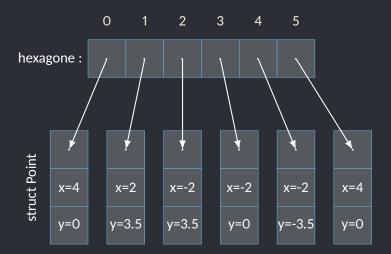
Profile:void free(void* ptr);

Exemple : free(tab);

Faites attention de ne pas libérer de mémoire qui n'est pas sur le tas.

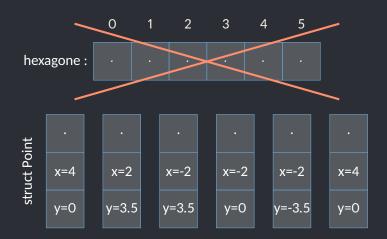
```
struct Point
 float x, y;
};
int main()
  struct Point** hexagone = malloc(6 * sizeof(struct Point*));
  for (int i = 0; i < 6; i++) {
    hexagone[i] = malloc(sizeof(struct Point));
   // Init of hexagone[i] ...
  }
  free(hexagone);
```

La mémoire : Fuitre mémoire



La mémoire : Fuitre mémoire

free(hexagone); // memory leak here !



La mémoire : Fuite mémoire

Explication

Ce que nous venons de voir est un exemple de fuite de mémoire. Une fuite de mémoire se produit lorsqu'une mémoire allouée dynamiquement n'est jamais libérée.

Pour résoudre le problème, nous devons procéder comme suit :

```
for (int i = 0; i < 6; i++) {
   free(hexagone[i]); // freach each point
}
free(hexagone); // free the array of pointers</pre>
```

Mémoire statique vs Mémoire dynamique

Avantages:

- Il n'y a pas de coût d'allocation.
- Cache local la plupart du temps car il est situé dans la pile.

Inconvénients:

- Très local, en raison de la nature de la pile.
- Taille limitée.
- La taille doit être fixée pendant la compilation ¹⁰.

^{10.} C autorise l'allocation de mémoire sur la pile dont la taille est déterminée lors 87/1 de l'exécution, ceci est interdit en C++.

Mémoire statique vs Mémoire dynamique

Avantages:

- Flexible, la taille peut être déterminée au moment de l'exécution.
- Globale.
- Peut gérer des tailles que la pile ne peut pas gérer.

Inconvénients:

- L'allocation peut être très coûteuse car elle nécessite un passage du mode utilisateur au mode noyau.
- Responsabilité de libérer la mémoire à la fin de l'utilisation.

malloc, calloc et realloc

malloc

```
Profile: void* malloc(size_t size);

Alloue ce qui lui est passé comme argument en octets mais n'effectue aucune initialisation.
```

Exemple:

malloc, calloc et realloc

calloc

```
Signature: void* calloc(size_t nmemb, size_t size);
Alloue nmemb * size octets et les initialise à zéro.
```

Exemple:

```
void* p1 = calloc(1, 256); // 512 bytes are allocated
int* p2 = (int*)calloc(4, sizeof(int)); // 4 * sizeof(int) bytes
    are allocated
struct A* p3 = (struct A*)calloc(2, sizeof(struct A)); // 2 *
        sizeof(struct A) bytes are allocated
assert(p2[0] == 0); // true
assert(p3[0].a == 0); // true
assert(p3[1].b == 0); // true
assert(p3[1].str[0] == 0); // true
```

malloc, calloc et realloc

realloc

Profile: void* realloc(void* ptr, size_t size);
Change la taille du bloc de mémoire pointé par ptr en size octets.

- Si size > taille de ptr: La mémoire pointée par le pointeur retourné par realloc sera de taille size. Le contenu de ptr est garanti d'être copié mais la mémoire ajoutée ne sera pas initialisée
- Si size < taille de ptr: Le contenu de ptr sera copié jusqu'à size octets, le reste du contenu de ptr sera ignoré. La taille de la mémoire pointée par la valeur de retour sera donc size.
- Si ptr est NULL : Cela aura le même effet que malloc.

malloc, calloc et realloc

realloc

• Si size est 0 et ptr n'est pas NULL : Cela aura le même effet que free.

N.B.:

- Sauf si ptr est NULL, il doit avoir été renvoyé par un appel antérieur à malloc(), calloc() ou realloc().
- L'accès et/ou l'écriture au pointeur a passé à realloc après l'appel est un comportement indéfini

malloc, calloc et realloc

Exemple

```
int* p1 = (int*)calloc(4, sizeof(int)); // 4 * sizeof(int) bytes
p[0] = 1; p[1] = 2; p[2] = 3; p[3] = 4;
int* p2 = (int*)realloc(p1, 6 * sizeof(int));
assert(p2[0] == 1); // true
assert(p2[1] == 2); // true
assert(p2[5] == 6); // Access to uninitialized memory !
assert(p1[5] == 6); // U.B !
```

malloc, calloc et realloc

Exemple

```
int* p3 = (int*)realloc(p2, 3 * sizeof(int));
assert(p3[0] == 1); // true
assert(p3[1] == 2); // true
assert(p3[3] == 3); // U.B!
assert(p2[3] == 3); // U.B!
int* p4 = (int*)realloc(p3, 0); // Equivalent to free(p3)
void* p5 = realloc(NULL, 8); // Equivalent to malloc(8)
void* p6 = realloc(NULL, 0); // Equivalent to malloc(0)
```

malloc, calloc et realloc

N.B.:

Le standard C ne dit rien quand 0 est passé à malloc (le comportement est spécifique au système d'exploitation).

- Sous Linux : malloc(0) renvoie NULL
- Sous Windows : malloc(0) renvoie un pointeur valide sur lequel free pourrait être appelé.

Les pointeurs

Définition

Un pointeur est une variable qui contient l'adresse d'une région de mémoire. Un pointeur peut contenir une adresse valide ou non (Exemple : le pointeur NULL).

Les pointeurs font généralement 4 ou 8 octets en fonction de l'architecture du processeur (32 ou 64 bits) :

- Avec un pointeur 32 bits, nous avons 4 Go de mémoire adressable.
- Avec un pointeur 64 bits, nous avons environ 17 milliards de Go de mémoire.

Les pointeurs

Syntaxes possibles

Toutes les syntaxes suivantes sont valides. Cependant, les deux premiers sont les plus courants :

```
Typename* ptr_name; // Recommended
Typename *ptr_name; // Recommended
Typename*ptr_name;
Typename * ptr_name;
```

Les pointeurs

Comment obtenir l'adresse d'une variable

Pour obtenir l'adresse d'une variable, l'opérateur & doit être utilisé.

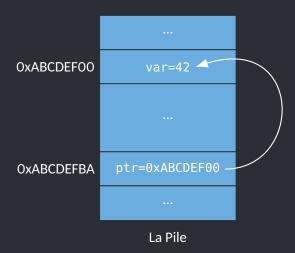
Exemple:

```
Typename var = ...;
Typename* ptr = &var;
```

Example

```
int var = 42;
int* ptr = &var;
```

Pointeur et variable sur la pile



Pointeur sur la pile et variable sur le tas



Tas 100/128

Les pointeurs

Attention:

Lorsque l'opérateur & est utilisé sur un tableau, il renvoie une adresse vers le premier élément de ce tableau et non un pointeur vers la première adresse du tableau.

Example :

```
int arr[] = {1, 2, 3};
int* ptrCpy = arr;
int* ptrArr = &arr;
assert(arr == &arr); // true
assert(ptrCpy == arr); // true
assert(ptrCpy == ptrArr); // true
assert(&ptrCpy == arr); // false
assert(&ptrCpy == &ptrArr); // false
```

Les pointeurs

Comment déréférencer un pointeur

Déréférencer un pointeur permet d'accéder à la valeur pointée par le pointeur. Pour cela, l'opérateur * doit être utilisé.

Example:

```
TypeName var = ...;
TypeName* ptr = &var;
assert(*ptr == var); // true
*ptr = ...; // changes the content of var
```

Les pointeurs

ATTENTION:

- Déréférencer un pointeur NULL est un comportement indéfini.
- Déréférencer un pointeur sur lequel free () a été appelé est un comportement indéfini.
- En général, le déréférencement d'un dangling pointer entraînera une corruption de pile ou un plantage du programme.

Définition

Un dangling pointer ¹¹ est un pointeur pointant vers une adresse mémoire valide qui a été libérée ou détruite.

Les pointeurs

```
Exemple 1:
int main() {
    int* ptr = NULL;
    {
        int x = 42;
        ptr = &x;
    }
    printf("content of ptr : %d\n", *ptr);
    *ptr = 3;
}
```

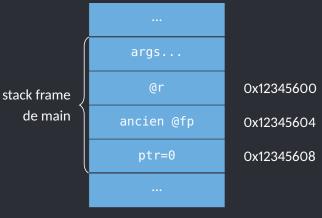
Dangling pointers

Solution 1:

```
int main() {
    int* ptr = NULL;
        int x = 42;
        ptr = &x;
    // x is out of scope (it's popped out of the stack)
    printf("content of ptr : %d\n", *ptr); // ptr is now a
                                            // dangling pointer
    *ptr = 3; // this will either corrupt the stack or
              // segfault
```

La mémoire : Dangling pointers

int* ptr = NULL;



La Pile

La mémoire : Dangling pointers

```
int x = 42;
ptr = &x;
```



La Pile

La mémoire : Dangling pointers

```
{
   int x = 42;
   ptr = &x;
} // x is out of scope (it's popped out of the stack)
```



Les pointeurs

Exemple 2: int main() { int* tab = malloc(4 * sizeof(int)); // Init tab to 1, 2, 3, 4 free(tab); tab[0] = 3; }

Les pointeurs : Dangling pointers

Solution 2: int main() { int* tab = malloc(4 * sizeof(int)); // Init tab to 1, 2, 3, 4 free(tab); // tab is now dangling tab[0] = 3; // will crash }

```
int* tab = malloc(4 * sizeof(int));
// Init tab to 1, 2, 3, 4
```



Tas 111/128

```
free(tab);
// tab is now dangling
```



Tas 112/128

Dangling pointers : Solution Possibles

Solution 1:

Une façon de résoudre ce problème consiste à affecter le pointeur à NULL après l'avoir libéré.

```
free(tab);
tab = NULL; // check if tab is NULL before we access it later
```

Problème avec la solution ci-dessus

Cette solution n'est pas parfaite si on libére tab dans une autre fonction et on l'affecte à NULL, seule la copie locale du pointeur sera NULL.

Dangling pointers : Solution Possibles

Solution 2:

Une meilleure façon de résoudre ce problème est de passer un pointeur vers le pointeur qui doit être libéré.

```
void myfree(void** pptr)
{
    if (pptr && *pptr) {
        free(*pptr);
        *pptr = NULL;
    }
}
```

Cette solution n'est pas parfaite non plus

Les pointeurs

Accès aux membres d'un struct/union

Il existe deux façons d'accéder à un champ struct via un pointeur :

```
struct StructName* ptr = &var;
ptr->field1_name = ...;
ptr->field2_name[0] = ...; // supposing that field2_name is an
         array this will change the first element of field2_name
ptr->field1_name; // accessing field1_name in the struct var
ptr->field2_name[0]; // accessing the first element in field2_name
         in the struct var
```

La mémoire Les pointeurs

Accès aux membres d'un struct/union

La deuxième façon est verbeuse et donc déconseillée :

```
struct StructName* ptr = &var;
(*ptr).field1_name = ...; // *ptr is between parentheses because
    "*" have a lower percedance level than "."
(*ptr).field2_name[0] = ...; // supposing that field2_name is an
    array this will change the first element of field2_name
(*ptr).field1_name; // accessing field1_name in the struct var
(*ptr).field2_name[0]; // accessing the first element in
    field2_name in the struct var
```

Arithmétique des pointeurs

Explication

Un pointeur n'est qu'une adresse mémoire. Cette adresse est une valeur numérique. Par conséquent, on peut effectuer des opérations arithmétiques sur un pointeur comme on peut le faire sur des valeurs numériques.

Il existe quatre opérateurs arithmétiques qui peuvent être utilisés sur les pointeurs : ++, -, + et -.

La formule

Pour un pointeur ptr avec le type Typename, l'expression ptr+step ajoutera step*sizeof(Typename) octets au pointeur ptr.

Arithmétique des pointeurs

Exemple

Lorsque vous utilisez l'arithmétique du pointeur, veillez à ne pas dépasser la taille allouée.

Pointeurs de fonction

Explication:

Comme les variables, les fonctions existent également en mémoire, nous pouvons donc avoir des pointeurs vers elles.

Cela peut être utile pour déterminer la fonction à exécuter au moment de l'exécution. On peut imiter le comportement des fonctions virtuelles fournies dans des langages tels que Java ou C++ en utilisant cette fonctionnalité.

Syntaxe:

ReturnType (*funcPtrName)(ArgType1, ArgType2, ..., ArgTypeN);

Pointeurs de fonction

Exemple

```
// includes ...
int max(int a, int b) { return a > b ? a : b; }
int min(int a, int b) { return a < b ? a : b; }</pre>
typedef int (*MyFuncType)(int, int); // MyFunc is a function
                                      // pointer type
int main() {
    int a = 5. b = 6:
    srand(time(NULL));
    int (*func)(int, int) = max; // Equivalent to
                                  // MyFuncType func = max;
    if (rand()%2) {
        func = &min; // same as func = main
    return func(a, b);
```

Pointer casting

// Pas trés important

Priorité des opérateurs

Table des opérateurs : 12

Priorité	Les opérateurs
1	++ () []>
2	++ + -! * & sizeof
3	* / %
4	+-
5	<< >>
6	<<= >>=
7	== !=
8	&

Table - Priorité des opérateurs en C

^{12.} pour plus de détails :

Priorité des opérateurs

Table des opérateurs : 13

Priorité	Les opérateurs
9	۸
10	1
11	&&
12	ll l
13	?:
14	= += -= *= /= %= <<= >>= &= ^= =
15	

Table - Priorité des opérateurs en C

^{13.} pour plus de détails :

Les Opérateurs et ordre d'évaluation

Pouvez-vous prédire la valeur de i sur chaque instruction sachant que i est initialement 0 :

```
i = (i++);  // i = 0 intially
i = ++i + i++;  // i = 0 intially
i = i++ + 1;  // i = 0 intially
f(++i, ++i);  // i = 0 intially
```

Les Opérateurs et ordre d'évaluation

```
i = ++i + i++; // undefined behavior i = i+++1; // undefined behavior f(++i, ++i); // undefined behavior f(i = -1, i = -1); // undefined behavior f(i, i++); // undefined behavior a[i] = i++; // undefined bevahior
```

Explication:

Pour une expression de type ExpA op ExpB ou op n'est pas & ou | | ou , l'ordre d'evaluation de ExpA et ExpB est indéfini (c'est-à-dire que ExpA peut être évalué avant ExpB ou vice versa)

Example

res = foo() + bar() foo peut être appelé avant bar ou vice versa 125/128

Les Opérateurs et ordre d'évaluation

Opérateurs qui garantissent l'ordre

Les opérateurs &&, $|\ |$ et , garantissent que l'expression à gauche est évaluée avant l'expression à droite.

L'opérateur ternaire (?:) garantit que la condition est évaluée en premier. Si la condition est vraie, l'expression de gauche à : est évaluée, sinon l'expression à droite de : est évaluée.

Example :

Chapitre 4

- Introduction
- Compilation
- 3. La langage C
- 4. Les outils

Conclusion