Programmation C

(Perfectionnement à la programmation en C)

Samuele Giraudo

Université Paris-Est Marne-la-Vallée LIGM, bureau 4B150 samuele.giraudo@u-pem.fr http://igm.univ-mlv.fr/~giraudo/

Introduction

L'objectif de ce module est d'approfondir les concepts de base et d'approcher certains concepts plus avancés de la programmation en C.

Celui-ci est organisé en trois axes.

1. Axe 1 : écrire un projet maintenable.

Bonnes habitudes de programmation, documentation, pré-assertions, gestion des erreurs, programmation modulaire, compilation.

2. Axe 2 : comprendre les mécanismes de base.

Pointeurs, allocation dynamique, types, types structurés, entrées/sorties.

3. Axe 3 : utiliser quelques techniques avancées.

Opérateurs bit à bit, mémoïsation, génération aléatoire, pointeurs de fonction, généricité.

Contenu du cours

Axe 1.	Axe 2.	Axe 3.
Bases	Allocation dynamique	Opérateurs
Habitudes	Entrées et sorties	Pointeurs de fonction
Modules	Types	Génération aléatoire
Compilation	Types structurés	Mémoïsation

Pré-requis

Ce cours demande les pré-requis suivants :

- des bases en programmation générale (notion de programme, d'instruction, de compilation);
- des bases en programmation impérative (notion de variable, de structures de contrôle, de fonction);
- ▶ des bases en **algorithmique** (manipulation de structures de données élémentaires comme les tableaux, les chaînes de caractères, les listes);
- ▶ des bases en **programmation en** C (écriture de fonctions, manipulation de tableaux, de chaînes de caractères, connaissance des types numériques, des types structurés, notion d'allocation dynamique, de pointeur).

Le langage C est apparu en 1972 dans les laboratoires Bell. Il a été créé par D. Ritchie et K. Thompson au même moment que l'apparition des systèmes UNIX.

Le langage C est apparu en 1972 dans les laboratoires Bell. Il a été créé par D. Ritchie et K. Thompson au même moment que l'apparition des systèmes UNIX.

Il est influencé par le langage B qui, contrairement au C, ne possédait pas de système de type.

Le langage C est apparu en 1972 dans les laboratoires Bell. Il a été créé par D. Ritchie et K. Thompson au même moment que l'apparition des systèmes UNIX.

Il est influencé par le langage B qui, contrairement au C, ne possédait pas de système de type.

Quelques dates sur l'évolution de la spécification du langage :

- ► 1978 : première description complète du langage (C traditionnel);
- ▶ 1989 : publication de la norme ANSI C (ou C89);
- ► 1995 : évolution du langage C94/95;
- ► 1999 : évolution du langage C99;
- ▶ 2011 : nouvelle version du standard C11.

Le langage C est apparu en 1972 dans les laboratoires Bell. Il a été créé par D. Ritchie et K. Thompson au même moment que l'apparition des systèmes UNIX.

Il est influencé par le langage B qui, contrairement au C, ne possédait pas de système de type.

Quelques dates sur l'évolution de la spécification du langage :

- ► 1978 : première description complète du langage (C traditionnel);
- ► 1989 : publication de la norme ANSI C (ou C89);
- ► 1995 : évolution du langage C94/95;
- ► 1999 : évolution du langage C99;
- ▶ 2011 : nouvelle version du standard C11.

Le C est un langage de bas niveau : il offre la possibilité de se placer « proche » de la machine. Il permet en effet de manipuler finement des données en mémoire, des adresses et des suites de bits.

Le C est un langage de bas niveau : il offre la possibilité de se placer « proche » de la machine. Il permet en effet de manipuler finement des données en mémoire, des adresses et des suites de bits.

Ce langage a été initialement pensé pour la conception de systèmes d'exploitation.

Il est cependant suffisamment expressif pour s'adapter à un large éventail d'utilisations.

Le C est un langage de bas niveau : il offre la possibilité de se placer « proche » de la machine. Il permet en effet de manipuler finement des données en mémoire, des adresses et des suites de bits.

Ce langage a été initialement pensé pour la conception de systèmes d'exploitation.

Il est cependant suffisamment expressif pour s'adapter à un large éventail d'utilisations.

Il se situe à la croisée des chemins du monde des langages de programmation : beaucoup de langages modernes sont traduits en C pour être compilés.

Le C est un langage de bas niveau : il offre la possibilité de se placer « proche » de la machine. Il permet en effet de manipuler finement des données en mémoire, des adresses et des suites de bits.

Ce langage a été initialement pensé pour la conception de systèmes d'exploitation.

Il est cependant suffisamment expressif pour s'adapter à un large éventail d'utilisations.

Il se situe à la croisée des chemins du monde des langages de programmation : beaucoup de langages modernes sont traduits en C pour être compilés.

L'un de ses compilateurs, gcc, fruit de nombreuses optimisations, fait que le C est un langage d'une efficacité extrême.

Axe 1 : écrire un projet maintenable

Bases

Habitudes

Modules

Compilation

Plan

Bases

Généralités

Expressions et instructions

Constructions syntaxiques

Variables

Fonctions et pile

Commandes préprocesseur

Plan

Bases

Généralités

Expressions et instructions

Constructions syntaxiques

Variables

Fonctions et pile

Commandes préprocesseur

La notion de programme

Un programme en C est avant tout un texte (contenu dans un **fichier**) qui suit certaines règles (dites de **syntaxe**).

La notion de programme

Un programme en C est avant tout un texte (contenu dans un **fichier**) qui suit certaines règles (dites de **syntaxe**).

C'est une collection de déclarations et de définitions de fonctions, de types, de variables globales, assorties de commandes pré-processeur.

La notion de programme

Un programme en C est avant tout un texte (contenu dans un **fichier**) qui suit certaines règles (dites de **syntaxe**).

C'est une collection de déclarations et de définitions de fonctions, de types, de variables globales, assorties de commandes pré-processeur.

Tout programme possède une **fonction principale** nommée main. C'est par elle que commence l'exécution du programme. On appelle ceci le point d'entrée du programme.

Compiler un programme signifie **traduire** le fichier le contenant en un langage compréhensible et exécutable par la machine cible.

Compiler un programme signifie **traduire** le fichier le contenant en un langage compréhensible et exécutable par la machine cible.

On compile un fichier Prog.c par la commande

gcc -o NOM Prog.c

Ceci produit un exécutable nommé NOM.

Compiler un programme signifie **traduire** le fichier le contenant en un langage compréhensible et exécutable par la machine cible.

On compile un fichier Prog.c par la commande

Ceci produit un exécutable nommé NOM.

On compilera obligatoirement avec les **options** -ansi, -pedantic et -Wall au moyen de la commande

```
gcc -o NOM -ansi -pedantic -Wall Prog.c
```

Compiler un programme signifie **traduire** le fichier le contenant en un langage compréhensible et exécutable par la machine cible.

On compile un fichier Prog.c par la commande

Ceci produit un exécutable nommé NOM.

On compilera obligatoirement avec les **options** -ansi, -pedantic et -Wall au moyen de la commande

► -ansi -pedantic : empêche un programme non compatible avec la norme ANSI C de compiler;

Compiler un programme signifie **traduire** le fichier le contenant en un langage compréhensible et exécutable par la machine cible.

On compile un fichier Prog.c par la commande

Ceci produit un exécutable nommé NOM.

On compilera obligatoirement avec les **options** -ansi, -pedantic et -Wall au moyen de la commande

```
gcc -o NOM -ansi -pedantic -Wall Prog.c
```

- -ansi -pedantic : empêche un programme non compatible avec la norme ANSI C de compiler;
- ► -Wall: active tous les messages d'avertissement.

Plan

Bases

Généralités

Expressions et instructions

Constructions syntaxiques

Variables

Fonctions et pile

Commandes préprocesseur

Une expression est définie récursivement comme étant soit

1. une constante;

Une expression est définie récursivement comme étant soit

- 1. une constante;
- 2. une variable;

Une expression est définie récursivement comme étant soit

- 1. une constante;
- 2. une variable;
- 3. une combinaison d'expressions et d'opérateurs;

Une expression est définie récursivement comme étant soit

- 1. une constante;
- 2. une variable;
- 3. une combinaison d'expressions et d'opérateurs;
- 4. un appel de fonction qui renvoie une valeur, c.-à-d., de type de retour autre que void.

Une expression est définie récursivement comme étant soit

- 1. une constante;
- 2. une variable;
- 3. une combinaison d'expressions et d'opérateurs;
- 4. un appel de fonction qui renvoie une valeur, c.-à-d., de type de retour autre que void.

Une expression n'est donc rien d'autre qu'un assemblage de symboles qui vérifie des règles syntaxiques et sémantiques.

Une expression est définie récursivement comme étant soit

- 1. une constante;
- 2. une variable;
- 3. une combinaison d'expressions et d'opérateurs;
- 4. un appel de fonction qui renvoie une valeur, c.-à-d., de type de retour autre que void.

Une expression n'est donc rien d'autre qu'un assemblage de symboles qui vérifie des règles syntaxiques et sémantiques.

Toute expression possède une valeur et un type. Le processus qui consiste à déterminer la valeur d'une expression se nomme l'évaluation.

P.ex., les entités suivantes sont des expressions :

Valeur: 0, type: int

```
Valeur: 0, type: int'a'Valeur: 97, type: int
```

```
Valeur: 0, type: int

'a'
Valeur: 97, type: int

"abc"
Valeur: "abc", type: char *
```

```
Valeur: 0, type: int
> 'a'
   Valeur: 97, type: int
► "abc"
   Valeur: "abc", type: char *
X
   Valeur : x, type : le type de la variable x
```

```
Valeur: 0, type: int
> 'a'
   Valeur: 97, type: int
▶ "abc"
   Valeur: "abc", type: char *
X
   Valeur : x, type : le type de la variable x
▶ a == 2
   Valeur: 0 ou 1, type: int
```

P.ex., les entités suivantes sont des expressions :

▶ 'a'
Valeur: 97, type: int

▶ "abc"
Valeur: "abc", type: char *

▶ xValeur : x, type : le type de la variable x

▶ a == 2
Valeur: 0 ou 1, type: int

P.ex., les entités suivantes sont des expressions :

Valeur: 0, type: int

> 'a'

Valeur: 97, type: int

Valeur: "abc", type: char *

▶ xValeur : x, type : le type de la variable x

▶ a == 2
Valeur: 0 ou 1, type: int

▶ b = 3
Valeur: 3, type: int

► f(5)

Valeur : la valeur renvoyée par f (5), type : le type de retour de f

P.ex., les entités suivantes sont des expressions :

0

Valeur: 0, type: int

> 'a'

Valeur: 97, type: int

▶ "abc"

Valeur: "abc", type: char *

X

Valeur : x, type : le type de la variable x

▶ a == 2

Valeur: 0 ou 1, type: int

b = 3

Valeur: 3, type: int

► f(5)

Valeur : la valeur renvoyée par f (5), type : le type de retour de f

 \triangleright (a == 0) && (b >= 3)

Valeur: 0 ou 1, type: int

P.ex., les entités suivantes sont des expressions :

Valeur: 0, type: int

> 'a'

Valeur: 97, type: int

▶ "abc"

Valeur: "abc", type: char *

▶ x

Valeur : x, type : le type de la variable x

▶ a == 2

Valeur: 0 ou 1, type: int

b = 3

Valeur: 3, type: int

► f(5)

Valeur : la valeur renvoyée par f (5), type : le type de retour de f

 \blacktriangleright (a == 0) && (b >= 3)

Valeur: 0 ou 1, type: int

▶ a > 'a' ? a : -a

Valeur : a ou -a, type : le type de a

P.ex., les entités suivantes sont des expressions :

0

Valeur: 0, type: int

> 'a'

Valeur: 97, type: int

▶ "abc"

Valeur: "abc", type: char *

▶ x

Valeur : x, type : le type de la variable x

▶ a == 2

Valeur: 0 ou 1, type: int

b = 3

Valeur: 3, type: int

► f(5)

Valeur : la valeur renvoyée par f (5), type : le type de retour de f

 \blacktriangleright (a == 0) && (b >= 3)

Valeur: 0 ou 1, type: int

▶ a > 'a' ? a : -a

Valeur : a ou -a, type : le type de a

1. + 7 * 2

Valeur: 15., type: float

Une instruction est définie récursivement comme étant soit :

- 1. une expression E; terminée par un point-virgule;
- 2. un bloc {I1 I2 ...Ik} où I1, I2, ..., Ik sont des instructions;
- 3. une conditionnelle if (E) I1 else I2, où E est une expression et I1 et I2 sont des instructions;
- 4. toute autre construction similaire (switch, while, do while, for);
- 5. une déclaration T X; de variable, où T est un type et X est un identificateur.

Une instruction est définie récursivement comme étant soit :

- 1. une expression E; terminée par un point-virgule;
- 2. un bloc {I1 I2 ...Ik} où I1, I2, ..., Ik sont des instructions;
- 3. une conditionnelle if (E) I1 else I2, où E est une expression et I1 et I2 sont des instructions;
- 4. toute autre construction similaire (switch, while, do while, for);
- 5. une déclaration T X; de variable, où T est un type et X est un identificateur.

À la différence des expressions, c'est souvent l'**effet** d'une instruction qui est sa raison d'être (et non plus uniquement sa valeur ou son type).

Une instruction est définie récursivement comme étant soit :

- 1. une expression E; terminée par un point-virgule;
- 2. un bloc {I1 I2 ...Ik} où I1, I2, ..., Ik sont des instructions;
- 3. une conditionnelle if (E) I1 else I2, où E est une expression et I1 et I2 sont des instructions;
- 4. toute autre construction similaire (switch, while, do while, for);
- 5. une déclaration T X; de variable, où T est un type et X est un identificateur.

À la différence des expressions, c'est souvent l'**effet** d'une instruction qui est sa raison d'être (et non plus uniquement sa valeur ou son type).

En effet, une instruction peut p.ex. afficher un élément, allouer une zone mémoire, lire dans un fichier, *etc.* La valeur de l'expression sous-jacente à l'instruction est d'importance secondaire.

P.ex., les entités suivantes sont des instructions :

▶ ; (instruction vide)

 \triangleright malloc(64);

1;

▶ int a;

 \triangleright a = b;

▶ int tab[64];

▶ while (1) a += 1;

ightharpoonup tab[3] = 9;

printf("abc\n");

► tab[3];

► a++;

return 31;

Une **expression** est à effet de bord (ou encore à effet secondaire) si son évaluation modifie la mémoire.

Une **expression** est à effet de bord (ou encore à effet secondaire) si son évaluation modifie la mémoire.

P.ex.,

▶ 0,▶ 'c',▶ tab[8],

sont des expressions qui ne sont pas à effet de bord.

Une **expression** est à effet de bord (ou encore à effet secondaire) si son évaluation modifie la mémoire.

P.ex.,

- **>** 0,
- · 'c',

- ▶ (a + 21) << 3,
- ► tab[8],

sont des expressions qui ne sont pas à effet de bord.

En revanche,

- \rightarrow a = 31,
- printf("abc\n"),

- char c,
- ightharpoonup malloc(64),

sont des expressions à effet de bord.

En règle générale, ce sont les éléments suivants dans les expressions qui produisent des effets de bord :

- les affectations;
- les allocations de mémoire;
- la sollicitation au système de fichiers.

En règle générale, ce sont les éléments suivants dans les expressions qui produisent des effets de bord :

- les affectations;
- les allocations de mémoire;
- la sollicitation au système de fichiers.

En revanche, les éléments suivants dans les expressions ne produisent pas d'effet de bord :

- lectures de constantes et de variables;
- calculs arithmétiques, logiques ou bit à bit;
- comparaisons.

Plan

Bases

Généralités

Expressions et instructions

Constructions syntaxiques

Variables

Fonctions et pile

Commandes préprocesseur

Un bloc est une suite d'instructions délimitée par des accolades. Il est constitué d'une partie consacrée à la déclaration de variables et d'une partie consacrée aux instructions.

Un bloc est une suite d'instructions délimitée par des accolades. Il est constitué d'une partie consacrée à la déclaration de variables et d'une partie consacrée aux instructions.

{ D I } lci, D est une section de **déclarations** et I est une section d'**instructions**.

Un bloc est une suite d'instructions délimitée par des accolades. Il est constitué d'une partie consacrée à la déclaration de variables et d'une partie consacrée aux instructions.

```
{
D
I
```

Ici, D est une section dedéclarations et I est une sectiond'instructions.

Les parties D et/ou I peuvent être vides.

Un bloc est une suite d'instructions délimitée par des accolades. Il est constitué d'une partie consacrée à la déclaration de variables et d'une partie consacrée aux instructions.

```
{
D
I
```

Ici, D est une section dedéclarations et I est une sectiond'instructions.

Les parties D et/ou I peuvent être vides.

Sachant qu'un bloc est une suite d'instructions, il est possible de placer un bloc dans la section d'instructions d'un bloc et ainsi d'**imbriquer** plusieurs blocs.

```
int main() {
    int b;
    scanf("%d", &b);
    {
        int a;
        a = 1 + b;
    }
}
```

La partie rouge est un bloc au sein d'une fonction main.

Il y a des règles concernant la visibilité des variables (que nous verrons plus loin).

```
int main() {
    int b;
    scanf("%d", &b);
        int a;
        a = 1 + b;
    int a;
    int b;
    scanf(" %d", &a);
        printf("%d\n", a);
    scanf(" %d", &b);
```

La partie rouge est un bloc au sein d'une fonction main.

Il y a des règles concernant la visibilité des variables (que nous verrons plus loin).

Ceci est un bloc constitué d'une section de déclarations et d'une section d'instructions.

Cette dernière partie contient un bloc (en rouge) dont la section de déclarations est vide.

Opérateur de test if

L'opérateur de test se décline en deux versions :

```
if (E)
B

if (E)
B_1
else
B_2
```

lci, E est une **expression booléenne**: elle est considérée
comme fausse si elle s'évalue en
zéro et comme vraie sinon. De plus,
B, B_1 et B_2 sont des blocs
d'instructions.

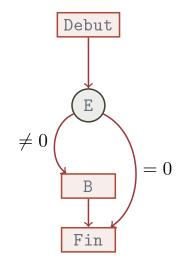
Opérateur de test if

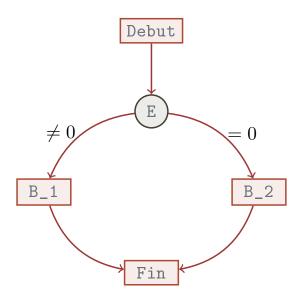
L'opérateur de test se décline en deux versions :

```
if (E)
B
if (E)
B_1
else
B_2
```

lci, E est une **expression booléenne**: elle est considérée
comme fausse si elle s'évalue en
zéro et comme vraie sinon. De plus,
B, B_1 et B_2 sont des blocs
d'instructions.

Diagrammes d'exécution :





L'instruction de branchement admet la syntaxe

```
switch (E) {
    case E_1 : I_1
    case E_2 : I_2
    ...
    case E_N : I_N
    default : I_D
}
```

lci, E, E_1, E_2, ..., E_N sont des expressions qui s'évaluent en des entiers.

Les I_1, I_2, ..., I_N et I_D sont des suites d'instructions.

L'instruction de branchement admet la syntaxe

```
switch (E) {
    case E_1 : I_1
    case E_2 : I_2
    ...
    case E_N : I_N
    default : I_D
}
```

lci, E, E_1, E_2, ..., E_N sont des expressions qui s'évaluent en des entiers.

Les I_1, I_2, ..., I_N et I_D sont des suites d'instructions.

L'exécution de cette instruction se passe ainsi. Soit i le plus petit entier (s'il existe) tel que les évaluations de E et E_i sont égales. Les instructions $I_i, ..., I_N$ ainsi que I_D sont exécutées.

L'instruction de branchement admet la syntaxe

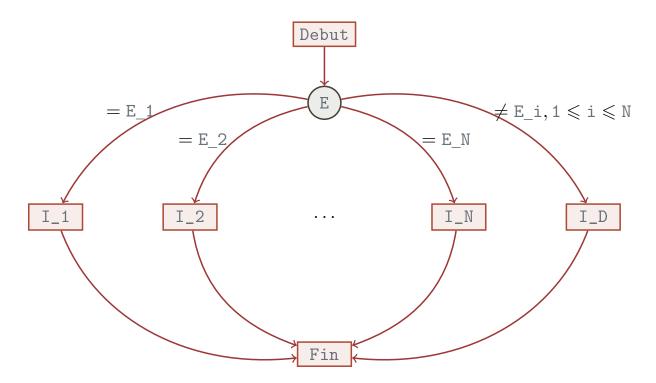
L'exécution de cette instruction se passe ainsi. Soit i le plus petit entier (s'il existe) tel que les évaluations de E et E_i sont égales. Les instructions $I_i, ..., I_N$ ainsi que I_D sont exécutées.

La ligne default : I_D est facultative. Si elle est présente, I_D est toujours exécutée.

On s'impose de terminer chaque I_j et I_D par le mot-clé break.

Ceci permet de n'exécuter que le I_i tel que les évaluations de E et E_i sont égales. Dans ce cas, I_D n'est exécuté que si aucun des E_i ne l'a été.

On obtient ainsi le diagramme d'exécution



Instruction itérative while

L'instruction itérative while admet la syntaxe

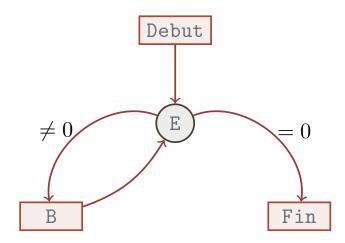
lci, E est une expression booléenne et B est un bloc d'instructions.

Instruction itérative while

L'instruction itérative while admet la syntaxe

lci, E est une expression booléenne et B est un bloc d'instructions.

Diagramme d'exécution :



Instruction itérative do while

L'instruction itérative do while admet la syntaxe

```
do
B
while (E);
```

lci, E est une expression booléenne et B est un bloc d'instructions.

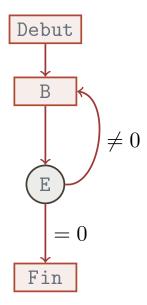
Instruction itérative do while

L'instruction itérative do while admet la syntaxe

```
do
B
while (E);
```

lci, E est une expression booléenne et B est un bloc d'instructions.

Diagramme d'exécution :



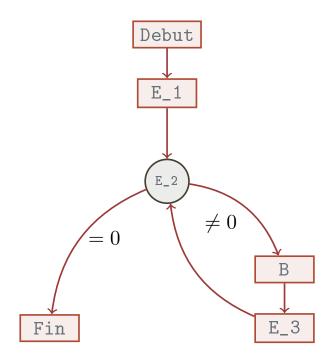
L'instruction itérative for admet la syntaxe

lci, E_2 est une expression booléenne (le test), E_1 (l'initialisation) et E_3
(l'incrémentation) sont des expressions et B est un bloc d'instructions.

L'instruction itérative for admet la syntaxe

lci, E_2 est une expression booléenne (le test), E_1 (l'initialisation) et E_3
(l'incrémentation) sont des expressions et B est un bloc d'instructions.

Diagramme d'exécution :



```
for (i = 0; i < 3; ++i) {
    printf("a");
}

for (x = lst; x != NULL;
    x = x->suiv) {
    afficher(x->elem);
}
```

Affiche trois occurrences du caractère 'a'.

Parcours une liste simplement chaînée et l'affiche. La variable x est un pointeur sur la cellule courante.

```
for (i = 0 ; i < 3 ; ++i) {
                                        Affiche trois occurrences du caractère
    printf("a");
                                         'a'.
                                        Parcours une liste simplement chaînée
for (x = 1st; x != NULL;
         x = x -> suiv) {
                                        et l'affiche. La variable x est un pointeur
                                        sur la cellule courante.
    afficher(x->elem);
}
for (; *str != '\0'; str++)) {
                                        Affiche la chaîne de caractères str. Il n'y a
    putchar(*str);
                                        pas d'initialisation. L'incrémentation fait
                                        pointer str sur le prochain caractère.
```

```
for (i = 0 ; i < 3 ; ++i) {
                                         Affiche trois occurrences du caractère
    printf("a");
                                         'a'.
                                         Parcours une liste simplement chaînée
for (x = 1st ; x != NULL ;
         x = x -> suiv) {
                                         et l'affiche. La variable x est un pointeur
                                         sur la cellule courante.
    afficher(x->elem);
}
for (; *str != '\0'; str++)) {
                                         Affiche la chaîne de caractères str. Il n'y a
    putchar(*str);
                                         pas d'initialisation. L'incrémentation fait
}
                                         pointer str sur le prochain caractère.
for (; *str != '\0';
                                         Affiche la chaîne de caractères str. Il
         putchar(*str++)) {
                                         n'y a pas d'initialisation. L'affichage est
    /* Rien. */
                                         réalisé comme effet de bord de
                                         l'incrémentation.
```

Les instructions de court-circuit sont break et continue.

Les instructions de court-circuit sont break et continue.

L'instruction break permet à l'exécution de sortir d'une structure switch, while, do while ou for. L'exécution continue alors aux instructions qui suivent cette structure.

Les instructions de court-circuit sont break et continue.

L'instruction break permet à l'exécution de sortir d'une structure switch, while, do while ou for. L'exécution continue alors aux instructions qui suivent cette structure.

L'instruction continue permet à l'exécution de sauter à la fin du bloc B d'une structure while, do while ou for. L'exécution continue alors à l'évaluation de l'expression test.

Dans une boucle for, l'instruction continue fait que l'expression d'incrémentation est tout de même évaluée.

Dans une boucle for, l'instruction continue fait que l'expression d'incrémentation est tout de même évaluée.

Dans une boucle for, l'instruction continue fait que l'expression d'incrémentation est tout de même évaluée.

Ce n'est pas le cas pour le break.

Dans une boucle for, l'instruction continue fait que l'expression d'incrémentation est tout de même évaluée.

Ce n'est pas le cas pour le break.