Utiliser des objets de type String

Maria-Virginia Aponte, François Barthélémy

NFA031

Dans ce chapitre, nous allons aborder l'utilisation de chaînes de caractères de type `String` et ce sera l'occasion de manipuler des objets Java. En effet les chaînes de caractères sont des objets. Nous allons apprendre à appeler des méthodes sur des objets et à créer de nouveaux objets de plusieurs façons.

1 Ce que l'on sait déjà de String

Il existe en java un type prédéfini pour les chaînes de caractères avec une syntaxe spéciale. Ce type s'appelle `String` et pour écrire une chaîne, on place les caractères de la chaîne entre guillemets. Il existe pour ce type un opérateur qui s'appelle la *concaténation* et qui s'écrit +. La concaténation permet de créer une chaîne en collant bout à bout deux chaînes existantes. Par exemple `"to"+"to"` crée la chaîne `"toto"`.

Par extension on peut concaténer une chaîne avec une valeur d'un type de base (int, boolean, double ou char). Il y a une conversion de type implicite et le résultat est une chaîne de caractère. Par exemple `"resultat: "+125` crée la chaîne : `"resultat: 125"`.

2 Naissance et vie des chaînes de caractères

Le type `String` existe dès l'origine dans le système java, mais ce n'est pas le cas des valeurs de ce type, c'est à dire des chaînes particulières que l'on peut utiliser. Par exemple, la chaîne `"bonjour"` n'existe pas au démarrage du système : si on veut l'utiliser il faut d'abord la créer. Pour cela, il existe quatre moyens :

- en écrivant cette chaîne avec ses guillemets dans le programme.
- au moyen de l'instruction `new`.
- au moyen de l'opérateur + : le résultat d'une concaténation est une nouvelle chaîne qui n'existait pas avant.
- au moyen d'une méthode qui renvoie une nouvelle chaîne.

Il existe plusieurs façons de créer une chaîne avec un `new`. Nous allons utiliser celle qui consiste à spécifier dans un tableau de caractères le contenu de la chaîne à créer.

```
char[] tab = {'b','o','n','j','o','u','r'};
String s = new String(tab);
```

Cela crée la chaîne `"bonjour"`. Si l'on n'a plus besoin du tableau après la création de la chaîne, on n'a pas besoin d'utiliser une variable pour désigner ce tableau, on peut écrire directement la création :

```
String s = new String(new char[]{'b','o','n','j','o','u','r'});
```

Une fois la chaîne créée, on peut l'afficher au moyen de `System.out.print` et l'utiliser comme opérande de l'opérateur de concaténation. On peut également appeler des méthodes.

Jusqu'à présent, les méthodes que nous avons écrites sont des méthodes des classes, déclarées avec le mot clé `static` et que l'on appelle avec le nom de la classe, un point et le nom de méthode

et les paramètres entre parenthèses. Par exemple dans `Math.random`, `Math` est le nom de la classe et `random` le nom de la méthode. Avec les chaînes de caractères, nous allons commencer à utiliser des méthodes des objets. Pour les appeler, il faut avant le point non pas un nom de classe, mais un objet, une valeur.

Prenons un premier exemple : la méthode `length()` renvoie la longueur de la chaîne. Elle ne prend pas de paramètre. La longueur est le nombre de caractères de la chaîne. Cette méthode n'a de sens que pour une chaîne donnée ; on l'appelle en mettant une chaîne ou un moyen d'en calculer une, avant le point et le nom de la méthode.

```
String s = "bonjour";
System.out.println(s.length());
```

L'appel `s.length()` va appeler la méthode `length` sur l'objet contenu dans la variable `s`. Le résultat est 7 puisqu'il y a 7 caractères dans `"bonjour"`.

On appelle souvent les méthodes en mettant un nom de variable avant le point, mais on peut mettre n'importe quelle expression qui calcule un objet du bon type. On peut appeler `length` non seulement sur une variable mais sur d'autres sortes d'expressions calculant un objet `String` :

— une chaîne entre guillement : `"bonjour".length()`
— une chaîne créér par new : `(new String()).length()`
— le résultat d'une concaténation : `(s + "qsd").length()`
— etc.

3 Les chaînes ressemblent un peu aux tableaux

Par certains côtés, les chaînes de caractères ressemblent aux tableaux plus qu'aux types de base tels que `int` ou `char`.

Voici les points communs entre tableaux et chaînes :

— il y a deux temps différents : la déclaration et la création d'une valeur.
— il y a possibilité de création explicite d'une nouvelle chaîne au moyen d'une instruction `new`.
— comme pour les tableaux, il y a possibilité de création implicite au moyen d'une syntaxe spéciale. Pour les tableaux, avec les accolades, pour les chaînes, avec les guillemets.
— ce sont des structures regroupant plusieurs valeurs dans un certain ordre.
— plusieurs noms différents peuvent être donnés à une même structure.
— les chaînes et les tableaux ont une longueur et cette longueur est supérieure ou égale à 0. Elle est fixe et invariable dans le temps.
— en revanche, une variable donnée peut contenir successivement des structures de tailles différentes.

Contrairement aux tableaux :

— il n'y a pas une syntaxe spécifique pour accéder directement aux caractères d'une chaîne de caractère.

Les chaînes de caractères sont le premier exemple d'objet que nous voyons en cours. Les types des objets sont comme les tableaux des types *références*, c'est à dire que les variables de ces types contiennent l'adresse des objets ou des tableaux en mémoire.

4 Quelques méthodes intéressantes

Voici quelques méthodes intéressantes :

— `charAt(int n)` : cette méthode renvoie le nième caractère de la chaîne, la numérotation commence à 0. Par exemple, si `s` est une `String`, `s.charAt(0)` renvoie le premier caractère (type `char`) de `s`.
— `toCharArray()` permet de transformer une chaîne en un tableau de `char`. Par exemple, si `s` est la `String "bonjour"` , `s.toCharArray()` renvoie un tableau de 7 `char` :

0	1	2	3	4	5	6
'b'	'o'	'n'	'j'	'o'	'u'	'r'

- `compareTo(String s2)` : compare deux chaînes selon l'ordre lexicographique (l'ordre du dictionnaire). Si `s1` et `s2` sont deux `String`, `s1.compareTo(s2)` renvoie un entier. Cet entier est négatif si `s1` est plus petit que `s2`, positif si `s1` est plus grand que `s2`, et 0 si `s1` et `s2` sont égales.
- `s1.toLowerCase()` et `s1.toUpperCase()` renvoient une nouvelle chaîne égale à `s1` mais avec toutes les lettres en minuscule et en majuscule respectivement.
- `trim()` : renvoie une chaîne dans laquelle les espaces en début et en fin de chaîne ont été supprimés. Par exemple `" truc chose ".trim()` renvoie la chaîne `"truc chose"`.
- `split(String s)` : découpe la chaîne en plusieurs morceaux en utilisant la chaîne `s` comme séparateur. Le résultat est un tableau de chaînes. Par exemple `"un;deux;trois".split(";")` renvoie le tableau `"un","deux","trois"`.
- `indexOf(String s)` : renvoie l'indice de la première occurrence de la chaîne `s` dans la chaîne. Par exemple `"un deux trois".indexOf("deux")` renvoie 3, car la chaîne `"deux"` commence à l'indice 3 de la chaîne `"un deux trois"`.
- `substring(int debut, int fin)` : renvoie la sous-chaîne de la chaîne sur laquelle la méthode est appelée comprise entre les indices `debut` et `fin`. Le caractère d'indice `debut` est inclus dans le résultat, mais pas celui d'indice `fin`. Par exemple `"bonjour".substring(2,4)` renvoie la sous-chaîne `"nj"`. Autrement dit, elle renvoie la sous-chaîne comprenant les caractères d'indice 2 et 3.

```
public class ExChaine2{
   public static void main (String [] arguments){
      String s1 = "bonjour";
      String s2;
      System.out.print("Entrez une chaine: ");
      s2 = Terminal.lireString();
      System.out.println(s1.charAt(0));
      System.out.println(s1.toUpperCase());
      System.out.println(s1.compareTo(s2));
      System.out.println("" + s1.equals(s2));
   }
}
```

A noter : il n'y a pas de méthode ni d'autre moyen de changer un caractère dans une chaîne : une fois la chaîne créée, on ne peut pas la modifier. Pour obtenir un résultat plus ou moins équivalent à un changement de caractère, il faut créer une nouvelle chaîne en concaténant des portions de la chaîne originale avec le caractère différent.

5 Variables et initialisations

Une valeur spéciale, `null`, est utilisable pour dire qu'une variable `String` ne pointe vers aucun objet. Ça n'est pas la même chose que *non initialisé*. Alors qu'une variable non initialisée ne peut pas être utilisée dans une expression parce qu'elle n'a pas de valeur, une variable qui contient la valeur `null` a une valeur et cette valeur peut par exemple être affichée. `System.out.println(uneVariable)` va afficher `null` si la variable contient cette valeur alors que le compilateur produira une erreur si la variable n'est pas initialisée.

Aucune méthode ne peut s'exécuter sur une variable qui contient `null`. Si l'on essaie de l'appeler cela provoque une erreur à la compilation ou à l'exécution.

Par exemple, dans le code suivant, `s` est non initialisée. le programme ne **compilatera** pas.

```
public class VNI{
   public static void main(String[] args){
      String s;
```

```
      System.out.println(s.length());
    }
}
```

Dans le code suivant, `s` est à `null`. Le programme compilera, mais on aura une erreur à l'**exécution :** une `NullPointerException` due à l'application de la méthode `length()` à une variable qui vaut `null`.

```
public class NPE {
   public static void main(String[] args){
      String s= null;
      System.out.println(s.length());
   }
}
```

Concrètement, `null` est utilisé pour dire qu'un objet n'est pas présent (imaginez par exemple qu'on représente une personne par un prénom, un second prénom, et un nom de famille, soit trois Strings. Si la personne n'a pas de second prénom, il pourra être initialisé à null).

Lorsque l'on crée un tableau de chaînes, la valeur `null` est la valeur par défaut placée dans toutes les cases. Dans l'état actuel de vos connaissance, c'est le cas où vous risquez le plus de rencontrer la valeur `null`.

Il est possible d'utiliser `null` comme valeur dans une affectation et de la tester dans un test d'égalité ou de différence. En revanche, on ne peut pas appeler de méthodes sur cette valeur. Par exemple `null` n'a pas de longueur, pas de premier caractère, etc.

```
public class VNI{
   public static void main(String[] args){
      String s = "truc";
      if (s!=null){
         System.out.println(s.length());
      }
      s = null;
      if (s!=null){
         System.out.println(s.length());
      }
   }
}
```

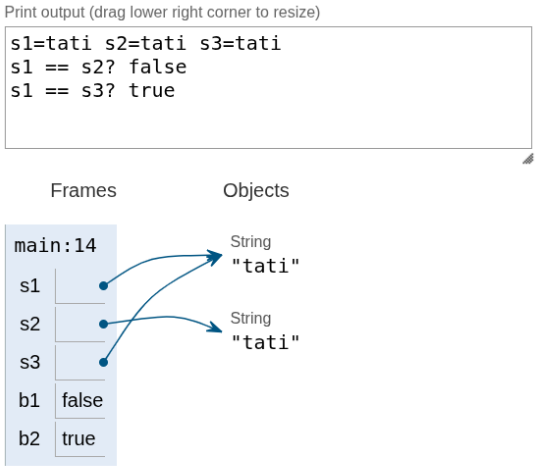
6 Comparaison de chaînes

Comme pour un tableau, les opérateurs `==` et `!=` ne regardent pas le contenu des chaînes mais juste leur adresse en mémoire. La question posée est : les deux chaînes sont-elles à la même adresse ? Ceci est illustré par le programme suivant.

```
public class EgCh{
   public static void main(String[] args){
      String s1, s2, s3;
      boolean b1, b2;
      s1 = "tati";
      s2 = "ta";
      s2 = s2 + "ti";
      s3 = s1;
      System.out.println("s1=" + s1 + " s2=" + s2 + " s3=" + s3);
      b1 = (s1 == s2);
      b2 = (s1 == s3);
```

```
        System.out.println("s1 == s2? " + b1);
        System.out.println("s1 == s3? " + b2);
    }
}
```

Si l'on représente par un dessin l'état de la mémoire après les affectations des trois chaînes, cela donne le dessin suivant produit par Pythontutor.



Si l'on veut réaliser une comparaison du contenu des chaînes et non plus de leur adresse, il faut utiliser une méthode. Pour une comparaison d'égalité, c'est la méthode `equals` qu'il faut appeler. Elle renvoie une valeur booléenne. Elle compare deux chaînes : l'une est l'objet sur lequel on appelle la méthode, l'autre est passée en paramètre. `s1.equals(s2)` fait un test d'égalité du contenu entre les deux chaînes `s1` et `s2`. Elles sont considérées comme égales si elles ont la même taille et les mêmes caractères dans le même ordre.

Le programme suivant illustre l'utilisation de `equals`. Notons au passage que l'on peut insérer un caractère guillement dans une chaîne de caractères en le faisant précéder du caractère `\` (voir la variable `s2`).n

```
public class ExChaine{
   public static void main(String[] args){
      String s1 = "Bonjour";
      String s2 = "C'est \"bien\" ";
      String s3;
      String[] ts = {"Paul", "Andre", "Jacques", "Odette"};
      System.out.println(s2);
      System.out.print("Entrez une chaine: ");
      s3=Terminal.lireString();
      System.out.println("s3: " + s3);
      s2 = "Bon";
      s3 = s2 + "jour";
      if (s1 != s3){
         System.out.println("Bizarre: s1 n'est pas egal a s3!");
         System.out.println("s1: " + s1 + ":");
         System.out.println("s3: " + s3+ ":");
      }
      if (s1.equals(s3)){
         System.out.println("s1 est quand meme egal a s3!");
      }
      if (!s1.equals(s3)){
```

```
            System.out.println("s1 n'est toujours pas egal a s3!");
        }
    }
}
```

La méthode `compareTo` fait une comparaison d'ordre selon le code unicode de chaque caractère. Cela correspond à l'ordre alphabétique uniquement pour les caractères sans accent. Elle renvoie un entier qui est 0 en cas d'égalité, un entier positif si l'objet est plus grand que la paramètre et un entier négatif si c'est le paramètre qui est plus grand.

7 Paramètre de la méthode main

Depuis le début de l'année, nous utilisons systématiquement la méthode `main` avec un paramètre de type `String[]`, c'est à dire un tableau de chaînes de caractères. Ce paramètre permet de transférer des informations entre la ligne de commande et le programme java. Prenons un exemple où le programme se contente d'afficher les valeurs passées sur la ligne de commande.

```
public class LigneCommande{
   public static void main(String[] args){
      for (int i=0; i < args.length; i++){
         System.out.println(args[i]);
      }
   }
}
```

Voici un exemple d'exécution :

```
> java LigneCommande un deux trois
un
deux
trois
```

La tableau `args` dans cette exécution a trois cases. Sa valeur est

0	1	2
¨un¨	¨deux¨	¨trois¨

Notons que même si l'on passe un nombre en paramètre, celui-ci est contenu dans le tableau sous forme d'une chaîne.

```
> java LigneCommande un 12 56 deux
un
12
56
deux
```

La tableaux args vaut

0	1	2	3
¨un¨	¨12¨	¨56¨	¨trois¨

Si l'on veut transformer une de ces chaînes en un entier, il faut utiliser une fonction de conversion.

8 Conversion entre chaînes et autres types

Il est parfois utile de convertir une chaîne de caractère en une valeur d'un autre type. Par exemple, on peut vouloir transformer une chaîne qui ne contient que des chiffres en un nombre entier. Pour réaliser la conversion, il faut utiliser la méthode `Integer.parseInt` et lui donner en paramètre la chaîne à convertir.

```
public class StringInt2{
   public static void main(String[] args){
      int x;
      String s = "12";
      x = Integer.parseInt(s);
      System.out.println(x);
   }
}
```

Pour convertir une valeur de type double, il faut utiliser la méthode `Double.parseDouble` et pour le type boolean, la méthode `Boolean.parseBoolean`.

Pour convertir dans l'autre sens, un int en chaîne, le plus simple est d'utiliser l'opérateur de concaténation : `""+12` (pour les doubles `""+12.3`, pour les booléens `""+true`). On concatène la chaîne vide avec la valeur à convertir.

9 Comparaison entre les chaînes et les tableaux de caractères

Les chaînes ressemblent à des tableaux de caractères, mais avec des différences qu'il convient de souligner.

Les ressemblances sont les suivantes :

— Ce sont deux structures contenant une séquence ordonnée de caractères
— Elles ont une longueur
— Les caractères sont numérotés au moyen d'indices entiers qui commencent à 0.
— Les chaînes comme les tableaux de `char` peuvent être affichés avec `System.out.print`.
— Les chaînes et les tableaux se parcourent avec une boucle `for` permettant d'examiner les caractères un à un.

Les différences sont les suivantes :

— Les tableaux de `char` n'ont aucune méthode permettant de les manipuler. Les `String` ont de nombreuses méthodes très utiles dont certaines ont été présentées dans ce document.
— Les caractères d'un tableau peuvent changer en affectant une nouvelle valeur à des cases de ce tableau alors que les caractère d'un `String` ne peuvent jamais changer. Ils sont fixés à la création de l'objet par `new` ou un équivalent de `new` et ne changent plus jamais par la suite pendant toute la durée d'exécution du programme. On dit que les chaînes sont *immuables* ou *non mutables*.

Conclusion

Les chaînes de caractères sont très utilisées. Ce premier exemple d'objet nous enseigne des choses que nous retrouverons avec tous les objets `Java` :

— Les objets n'existent pas si on ne les crée pas
— Une façon de créer un nouvel objet est d'utiliser `new`
— Une autre façon de créer un nouvel objet est d'invoquer une méthode qui renvoie un nouvel objet
— Les objets ont des méthodes que l'on peut appeler avec la notation pointée.
— La valeur `null` peut être utilisée comme valeur pour une variable qui ne contient pas d'objet.