Les Pièges du Go 7

Michel Casabianca
casa@sweetohm.net

Le langage de programmation Go a la réputation d'être simple à apprendre. Cependant, il recèle quelques pièges qui peuvent être difficiles à détecter pour un novice. Cette série d'articles propose d'en désamorcer quelques uns.

Nous voulons afficher la longueur d'une chaîne de caractères. Nous écrivons donc le code suivant :

```
package main

func main() {
        s := "échec"
        println("length of", s, "is", len(s))
}
```

Sur le Playground

Mais lorsque nous exécutons ce code, nous obtenons :

```
$ go run broken.go
length of échec is 6
```

Ce n'est probablement pas ce à quoi nous nous attendions... Pourquoi ? Comment corriger ce code ?

Explication

L'appellation *chaîne de caractères* est inexacte dans le cas du Go. Nous devrions plutôt parler de tableau de `bytes`. Pour être exact, une chaîne de caractère du Go est un tableau des octets de ladite chaîne encodée en *UTF-8*. Or notre chaîne comporte un caractère accentué, qui n'est pas encodé en *UTF-8* sur un octet, mais sur 2. Par conséquent la taille du tableau d'octets qui encode cette chaîne est *6* et non *5*.

Reste à savoir maintenant comment déterminer le nombre de caractères (ou de *Runes*) dans cette chaîne...

La première solution consiste à convertir la chaîne en un tableau de *Runes* dont on peut obtenir la tailles :

```
package main

func main() {
        s := "échec"
```

```
        println("length of", s, "is", len([]rune(s)))
}
```

Sur le Playground

Une autre solution consiste à utiliser la fonction `RuneCountInString()` du package *utf8* :

```
package main

import "unicode/utf8"

func main() {
        s := "échec"
        println("length of", s, "is", utf8.RuneCountInString(s))
}
```

Sur le playground

Dans ces deux cas, nous obtenons le bonne taille :

```
$ go run fixed.go
length of échec is 5
```

Ce qu'il est important de retenir c'est que `len(string)` ne donne la longueur d'une chaîne que dans le cas où celle-ci ne comporte que des caractères ASCII. Il ne faut donc **pas utiliser la fonction `len(string)` dans le cas général**.

Enjoy!