Working with Strings

Steve Gordon

.NET Engineer and Microsoft MVP

@stevejgordon www.stevejgordon.co.uk

Overview

Learn about Globalization

Configure application culture

Validate string data

Escape special characters in strings

Create verbatim string literals

Globalization

Countries have different rules governing formatting and handling of data.

$500.00

Formatting Differences

United Kingdom	Germany
2,125,000.00	2.125.000,00
£500.00	500,00€

Rules can change over time as new standards are adopted

Applications which parse or display data must respect local preferences

- Known as globalization

.NET relies on the OS for cultural rules

Linux uses Internal Components for Unicode (ICU)

Prior to Windows 10 (May 2019 update) Windows used National Language Support (NLS)

.NET 5 and later prefer ICU on Windows when available

Globalization is crucial to ensure applications apply the correct rules when parsing, formatting and sorting string data.

Culture is particularly important when presenting text or working with strings which may contain culture-specific formatting

The culture of a .NET application defaults to that of the runtime OS

Threads can be assigned a specific culture at runtime

Culture and globalization are particularly important when presenting text or working with strings which may contain culture-specific formatting.

Demo

Set application culture

Create a string literal

String literals are treated in a special way by the Roslyn compiler and .NET runtime.

String Interning

The .NET CLR maintains a table called the "Intern Pool", stored in the large object heap

- Used to deduplicate strings

Unique string literals are stored in the intern pool

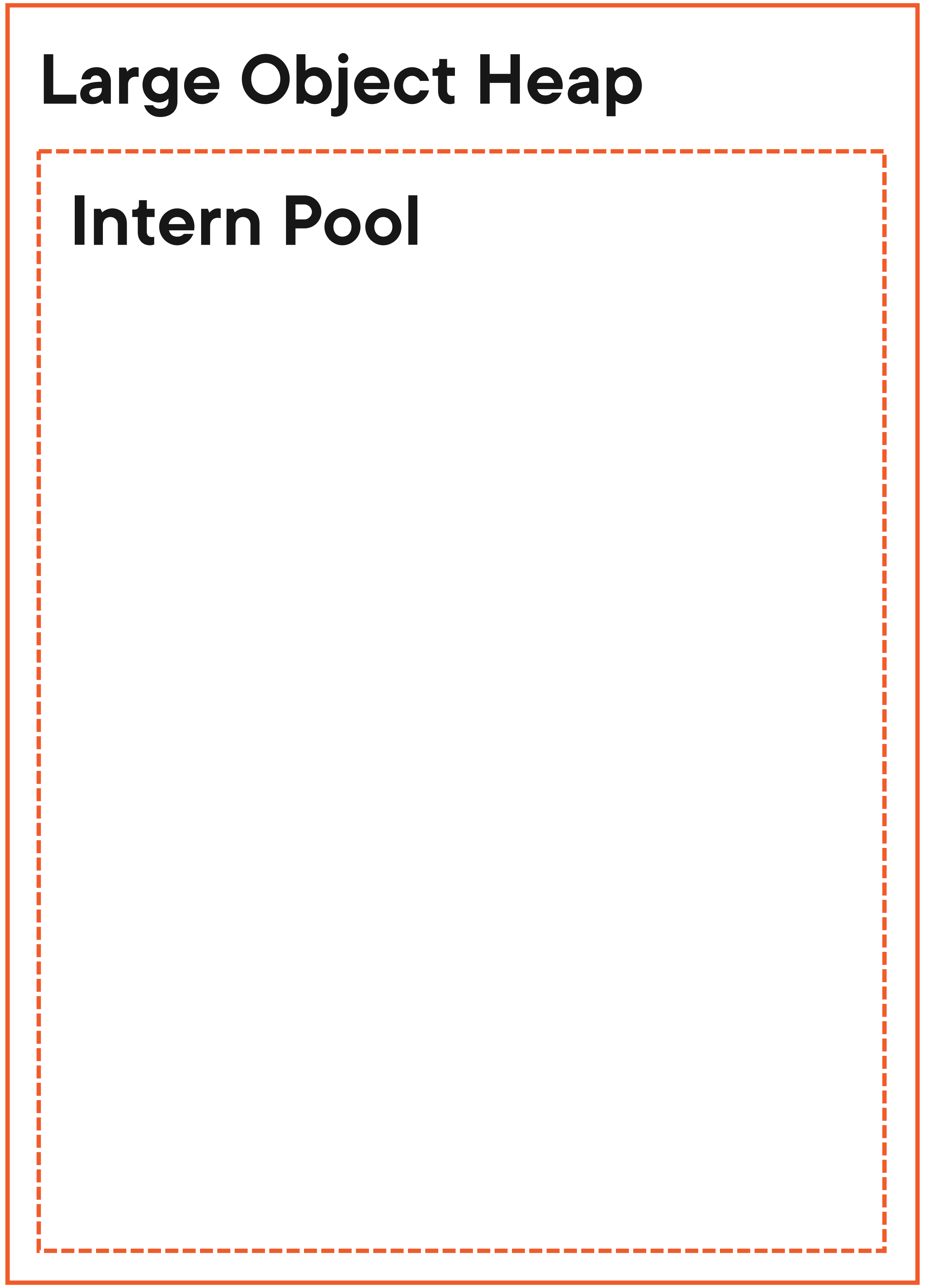
Avoids repeat allocations for the same string value

```
...

var firstString = “Hello, world”

...

var anotherString = Hello, world”
```

Large Object Heap

Intern Pool

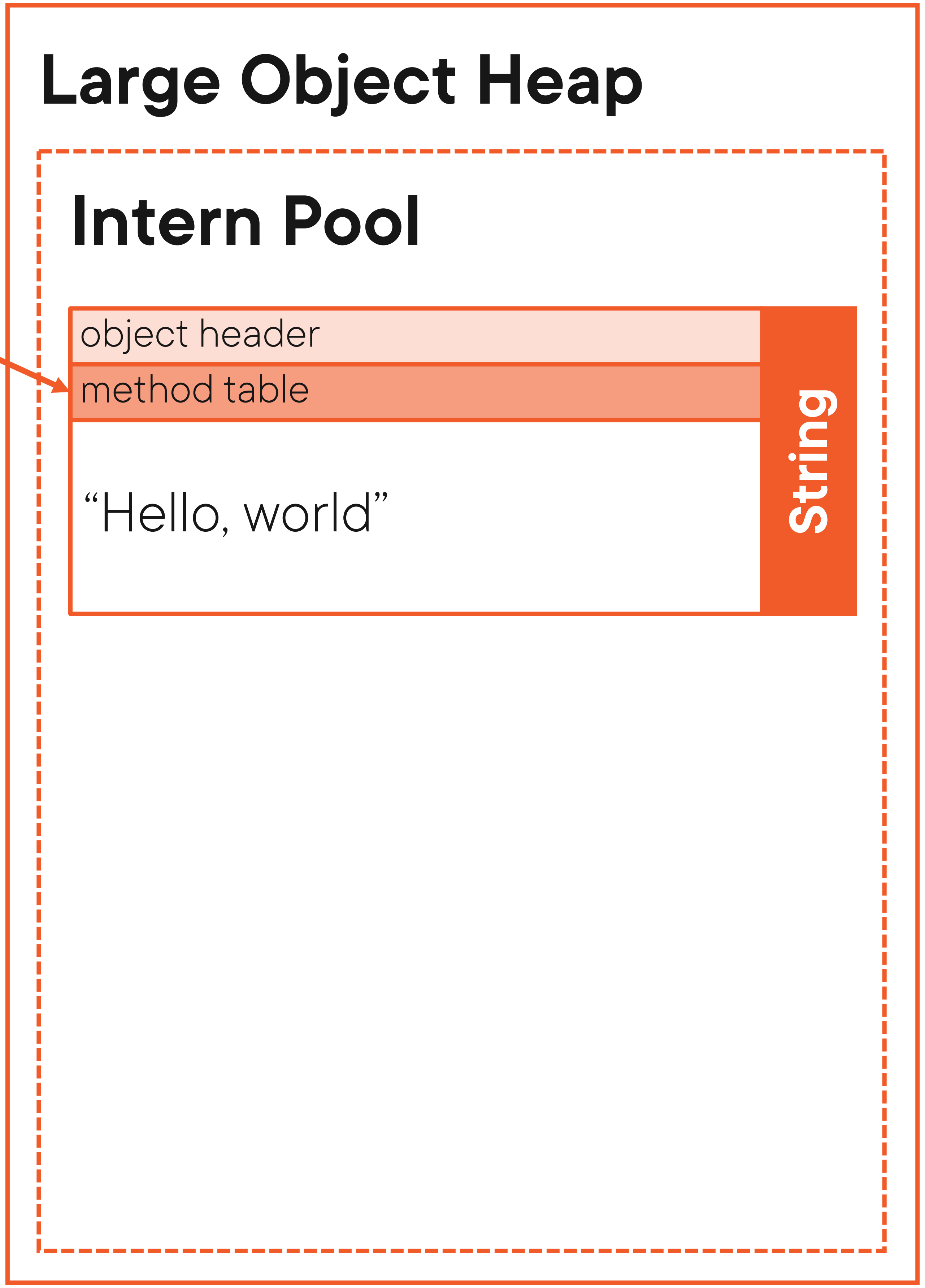
```
...

var firstString = “Hello, world”

...

var anotherString = Hello, world”
```



Large Object Heap

Intern Pool

```
...
var firstString = “Hello, world”

...

var anotherString = Hello, world”
```



Large Object Heap

Intern Pool

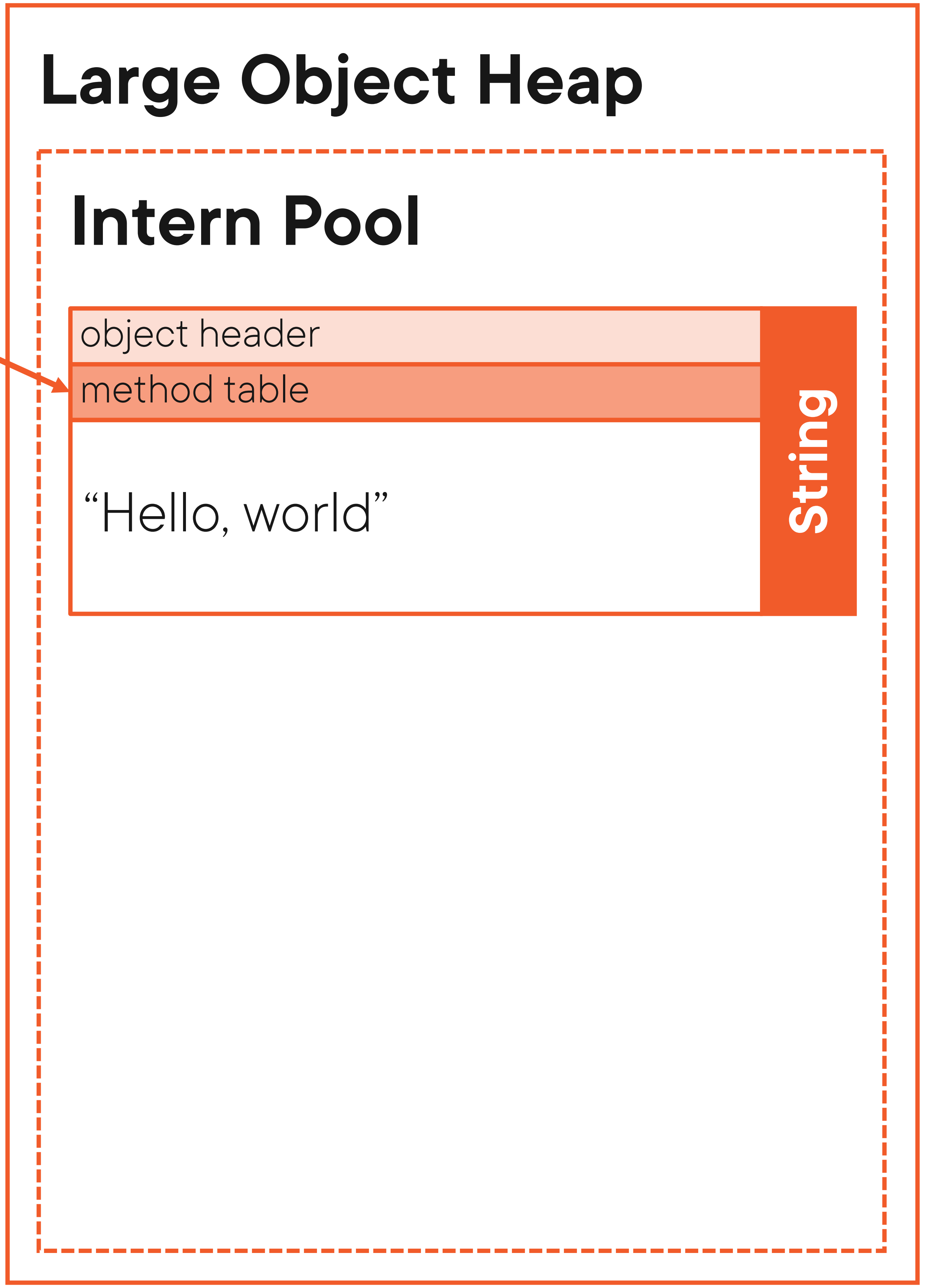
object header

method table

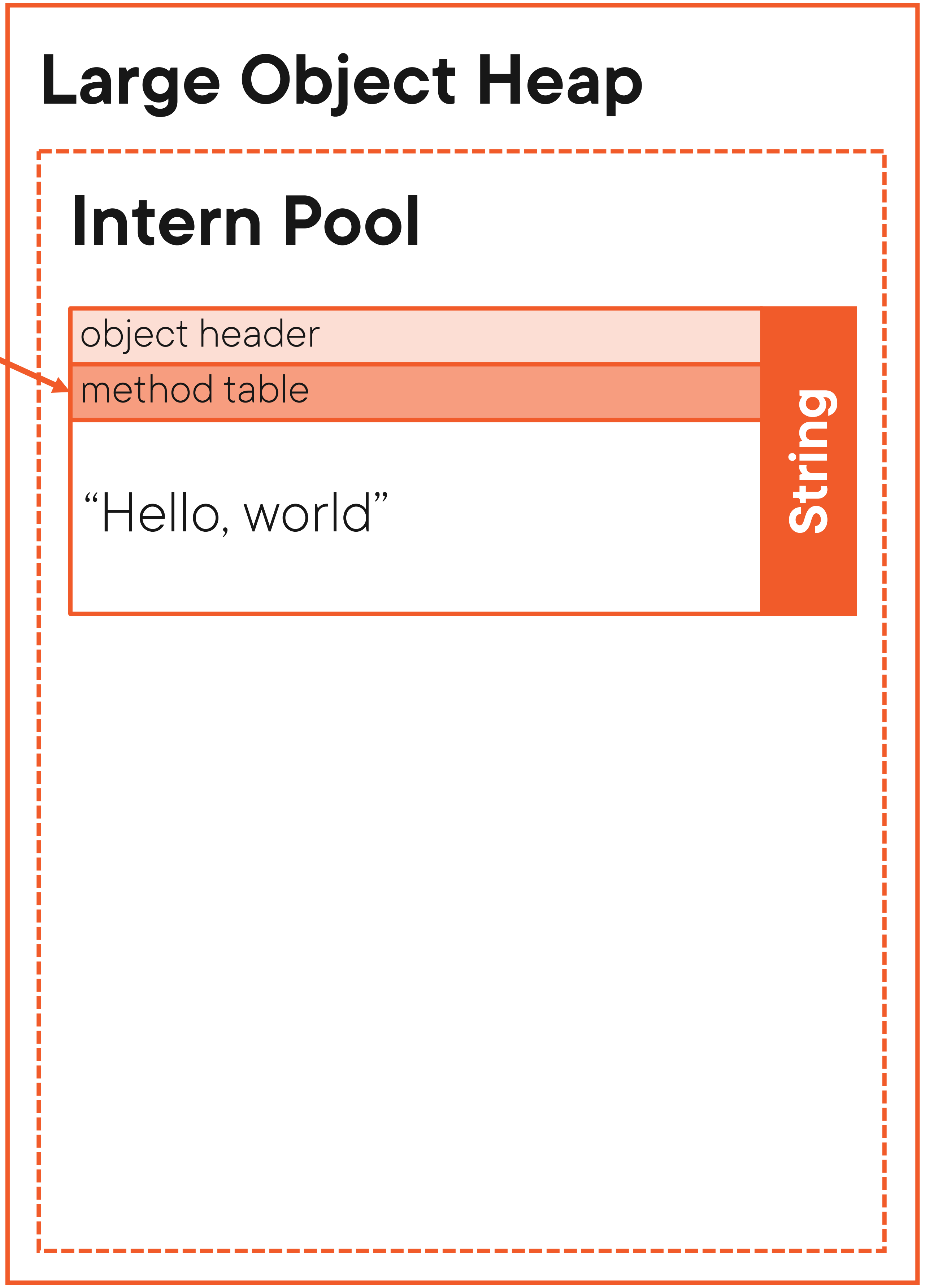
“Hello, world”

String

```
...

var firstString = “Hello, world”

...

var anotherString = Hello, world”
```



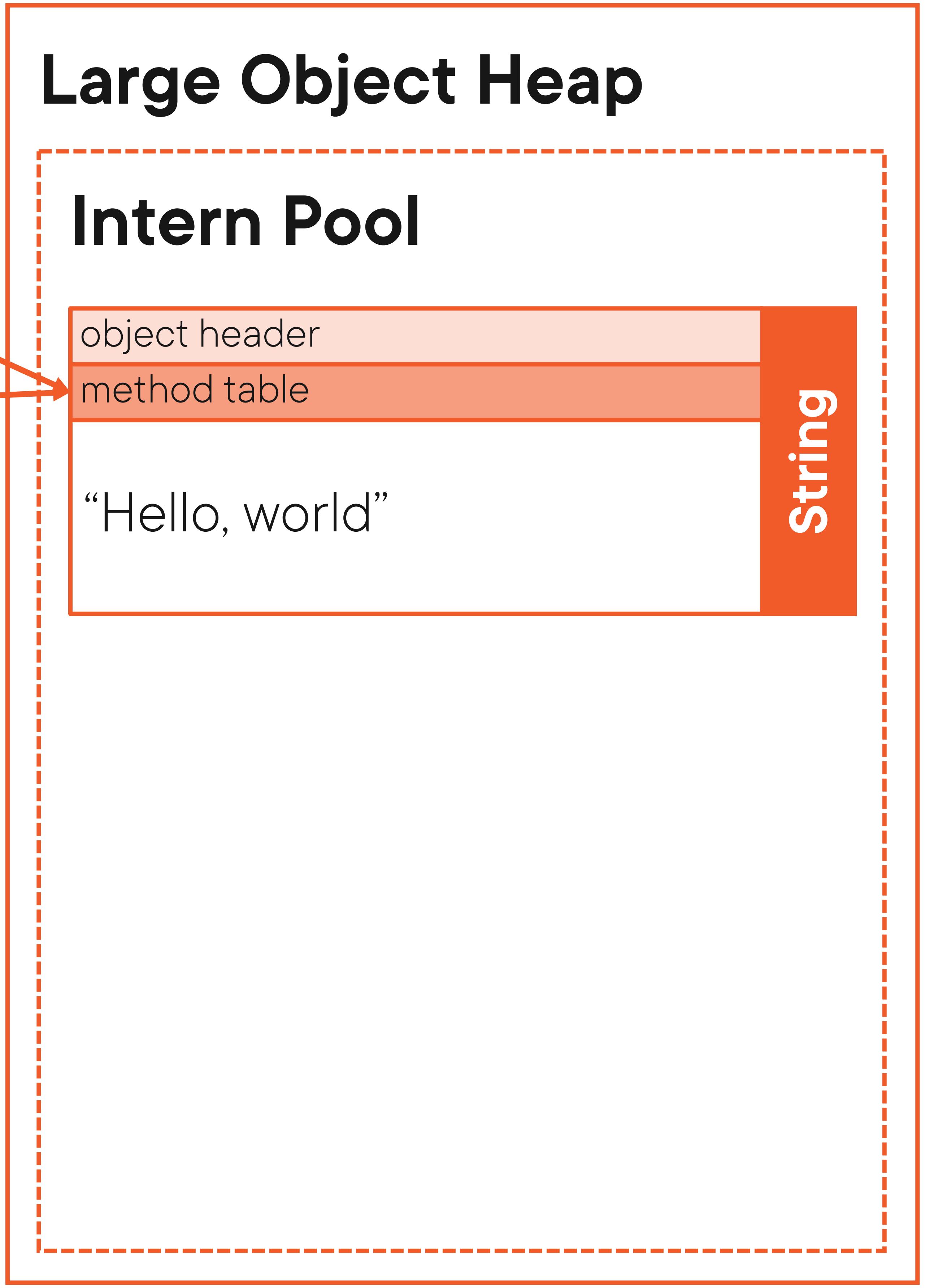
```
...

var firstString = “Hello, world”

...

var anotherString = Hello, world”
```



Large Object Heap

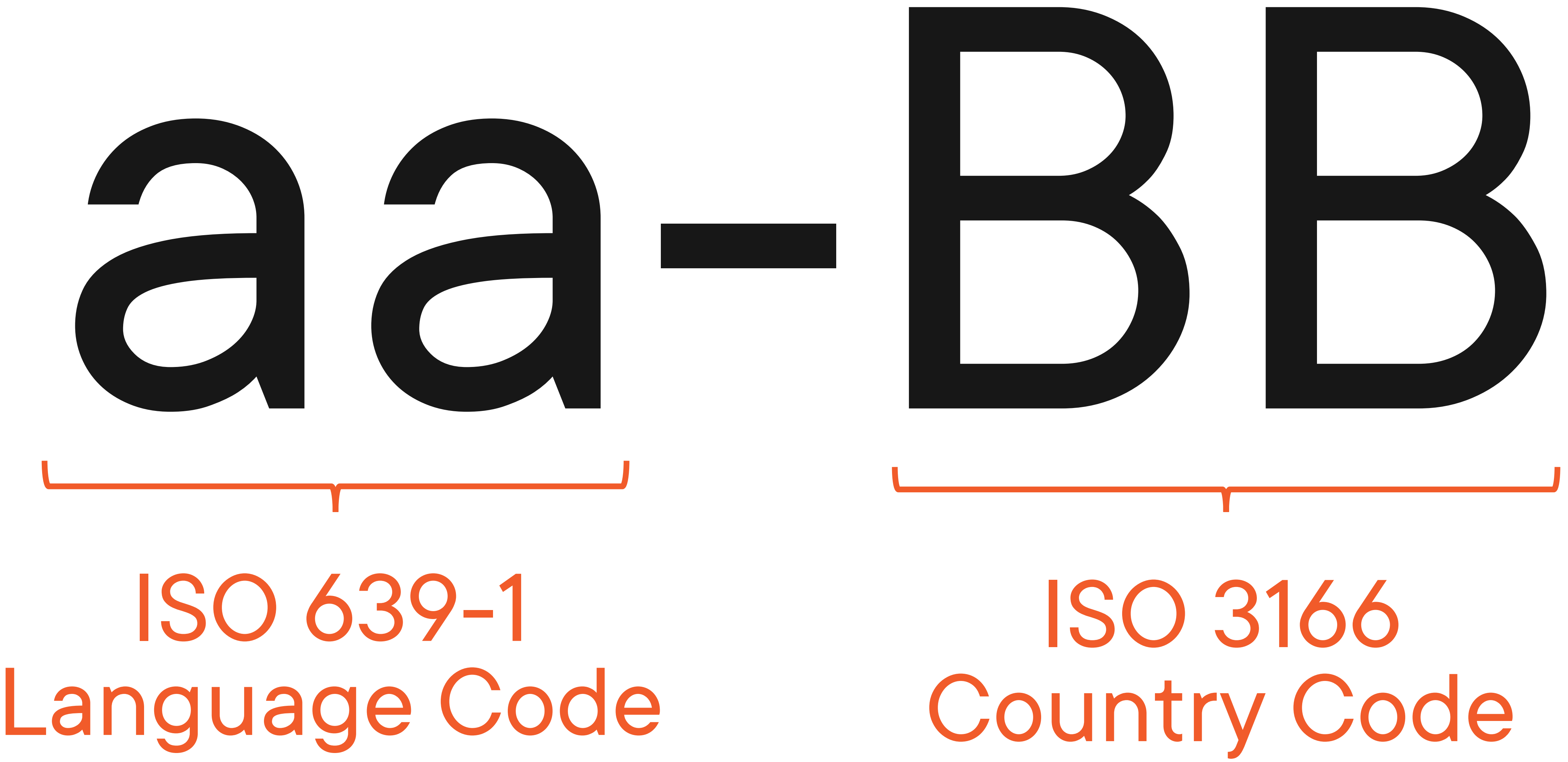
Intern Pool

```
...

var firstString = “Hello, world”

...

var anotherString = Hello, world”
```



Strings created at runtime may be manually interned, but this is an advanced topic and rarely needed.

CultureInfo Name



aa-BB

ISO 639-1
Language Code

ISO 3166
Country Code

Demo

Accept string data

Validate string data

Requirements

- Collect basic user information from the user.
    - Forename
    - Surname
    - Department ID

String Convenience Methods

IsNullOrEmpty
Indicates whether the specified string is null or an empty string ("").

IsNullOrWhitespace
Indicates whether a specified string is null, empty, or consists only of white-space characters.

```
var a = "   ";
Console.WriteLine(string.IsNullOrEmpty(a));             ◀ False
Console.WriteLine(string.IsNullOrWhiteSpace(a));        ◀ True


a = "";
Console.WriteLine(string.IsNullOrEmpty(a));             ◀ True
Console.WriteLine(string.IsNullOrWhiteSpace(a));        ◀ True


a = null;
Console.WriteLine(string.IsNullOrEmpty(a));             ◀ True
Console.WriteLine(string.IsNullOrWhiteSpace(a));        ◀ True
```

Demo

Escape special characters in strings

Requirements

- Include an introductory message in the console, including the application code name.
- Include an instruction message before accepting user data.
- Set a valid default directory in ProcessingOptions.

Requirement

- Include a face with tears of joy emoji at the end of the introductory message.

Demo

Use verbatim string literals

Verbatim String Literals

Use a mode that indicates to the compiler that all characters are to be interpreted literally.

Up Next: Processing and Parsing Strings