

# Programación 1

## Tema 10

---

### Representación de cadenas de caracteres



Escuela de  
Ingeniería y Arquitectura  
**Universidad Zaragoza**





# Índice

---

- ❑ Caracteres
- ❑ Cadenas de caracteres
- ❑ Problemas

# El tipo carácter

---

- Tipos carácter
  - **Dominio** de valores
  - **Representación** de los valores
    - Externa (en C++)
    - Interna (en la memoria del computador)
  - **Operadores** asociados

# Caracteres

## Dominio de valores

---





# Caracteres

## Dominio de valores

---

- ❑ Letras mayúsculas del alfabeto inglés
- ❑ Letras minúsculas del alfabeto inglés
- ❑ Dígitos
- ❑ Signos de puntuación
- ❑ Signos matemáticos
- ❑ Letras con diacríticos (alfabetos latinos occidentales)
- ❑ Letras alfabetos centro-europeos
- ❑ Letras alfabeto griego
- ❑ Letras alfabeto cirílico
- ❑ Letras alfabetos asiáticos

# Caracteres

## Unicode



- Estándar de codificación de caracteres
- Dominio de valores:
  - Alfabeto latino: A b c d e f g h i j k l m n o p q r s t u v w x y z
  - Alfabeto griego: α β γ δ ε ζ η θ ι κ λ μ ν ξ ο π ρ ς σ τ υ φ χ ω
  - Alfabeto cirílico: б в г ж з и й к л м н п с т у ф х ц ч ш щ ъ ы ь
  - Alfabetos centro-europeos: Á â ã ä Í Ć ç Č é ě ě ě í î ð ð ñ ñ ó ô ř ř
  - Alfabeto árabe: ی و م ل ع ص س د خ ح ج ث ت ة ب ا ی ا ؤ ا ء ک گ ژ چ
  - Alfabeto hebreo: ת ש ר ק צ ץ פ ף ע ס נ ן מ ם ל כ ך י ט ח ז ה ד ג ב א
  - Alfabetos asiáticos: 中文萬國碼際字出典フリ百科事典ィキペデア
  - Símbolos: £ Pts € № ¼ ½ ¾ ⅓ ← ↑ ↗ ⇔ ∇ ∂ ∃ ∄ ∅ ∞ ∫
  - Emoji: 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊

# Caracteres en C++

---

- Dos tipos
  - **char**
    - 1 *byte* (8 bits)
  - **wchar\_t**
    - 2 *bytes* (16 bits) en GNU GCC

# Caracteres

- **char**
- **Dominio de valores**
  - 95 caracteres
    - Letras del alfabeto inglés
    - Dígitos
    - Signos de puntuación
    - Otros símbolos
  - 33 *caracteres* de control

|    |   |   |   |   |   |
|----|---|---|---|---|---|
|    | 0 | @ | P | ` | p |
| !  | 1 | A | Q | a | q |
| "  | 2 | B | R | b | r |
| #  | 3 | C | S | c | s |
| \$ | 4 | D | T | d | t |
| %  | 5 | E | U | e | u |
| &  | 6 | F | V | f | v |
| '  | 7 | G | W | g | w |
| (  | 8 | H | X | h | x |
| )  | 9 | I | Y | i | y |
| *  | : | J | Z | j | z |
| +  | ; | K | [ | k | { |
| ,  | < | L | \ | l |   |
| -  | = | M | ] | m | } |
| .  | > | N | ^ | n | ~ |
| /  | ? | O | _ | o |   |



# Caracteres

## □ Representación externa en C++

- 'a' 'A' 'b' 'B' 'z' 'Z'
- '0' '1' '2' '3' '4' '5' '6' '7'  
'8' '9'
- '+' '-' '\*' '/' '<' '=' '>'
- '(' ')' '[' ']' '{' '}'
- '#' '\$' '%' '&' ',' '.' ':' ';'   
 '!' '?' '@' '^' \_ ` ' | ' ~ '  
'\_'
- '"' '\ ' '\\'

# Representación interna

---

- Codificación arbitraria en binario
  - Código ASCII
    - *American Standard Code for Information Interchange*
    - Estandarizada por la American Standards Association en 1963
- Ejemplo: 'A' se codifica con
  - la secuencia binaria 0100 0001
  - el código numérico 65

# Representación interna

| Código | Carácter | Código | Carácter | Código | Carácter | Código | Carácter | Código | Carácter | Código | Carácter | Código | Carácter | Código | Carácter | Código | Carácter |
|--------|----------|--------|----------|--------|----------|--------|----------|--------|----------|--------|----------|--------|----------|--------|----------|--------|----------|
| 0      | NUL      | 16     | DLE      | 32     |          | 48     | 0        | 64     | @        | 80     | P        | 96     | `        | 112    | p        |        |          |
| 1      | SOH      | 17     | DC1      | 33     | !        | 49     | 1        | 65     | A        | 81     | Q        | 97     | a        | 113    | q        |        |          |
| 2      | STX      | 18     | DC2      | 34     | "        | 50     | 2        | 66     | B        | 82     | R        | 98     | b        | 114    | r        |        |          |
| 3      | ETX      | 19     | DC3      | 35     | #        | 51     | 3        | 67     | C        | 83     | S        | 99     | c        | 115    | s        |        |          |
| 4      | EOT      | 20     | DC4      | 36     | \$       | 52     | 4        | 68     | D        | 84     | T        | 100    | d        | 116    | t        |        |          |
| 5      | ENQ      | 21     | NAK      | 37     | %        | 53     | 5        | 69     | E        | 85     | U        | 101    | e        | 117    | u        |        |          |
| 6      | ACK      | 22     | SYN      | 38     | &        | 54     | 6        | 70     | F        | 86     | V        | 102    | f        | 118    | v        |        |          |
| 7      | BEL      | 23     | ETB      | 39     | '        | 55     | 7        | 71     | G        | 87     | W        | 103    | g        | 119    | w        |        |          |
| 8      | BS       | 24     | CAN      | 40     | (        | 56     | 8        | 72     | H        | 88     | X        | 104    | h        | 120    | x        |        |          |
| 9      | HT       | 25     | EM       | 41     | )        | 57     | 9        | 73     | I        | 89     | Y        | 105    | i        | 121    | y        |        |          |
| 10     | LF       | 26     | SUB      | 42     | *        | 58     | :        | 74     | J        | 90     | Z        | 106    | j        | 122    | z        |        |          |
| 11     | VT       | 27     | ESC      | 43     | +        | 59     | ;        | 75     | K        | 91     | [        | 107    | k        | 123    | {        |        |          |
| 12     | FF       | 28     | FS       | 44     | ,        | 60     | <        | 76     | L        | 92     | \        | 108    | l        | 124    |          |        |          |
| 13     | CR       | 29     | GS       | 45     | -        | 61     | =        | 77     | M        | 93     | ]        | 109    | m        | 125    | }        |        |          |
| 14     | SO       | 30     | RS       | 46     | .        | 62     | >        | 78     | N        | 94     | ^        | 110    | n        | 126    | ~        |        |          |
| 15     | SI       | 31     | US       | 47     | /        | 63     | ?        | 79     | O        | 95     | _        | 111    | o        | 127    | DEL      |        |          |

# Otras codificaciones de caracteres

---

- 8 bits → 256 caracteres
  - Latin1 (ISO 8859-1), Latin0 (ISO 8859-15), Windows-1252
  - Página de códigos 850
- 16 bits → 65 536 caracteres
  - UCS-2 (2-byte Universal Character Set)
- Longitud variable → ~137 000 caracteres definidos por Unicode
  - UTF-8
  - UTF-16



# Universal Character Set (UCS)

---

- Estándar internacional ISO/IEC 10646 (~Unicode)
  - Define 110 000 **caracteres abstractos**
  - Cada carácter abstracto se identifica de forma precisa por un entero único: **punto de código** (*code point*)
  - Cada punto de código se puede codificar de acuerdo con distintas **codificaciones**:
    - UTF-8
      - 1, 2, 3 o 4 bytes
      - Compatible con los códigos ASCII de 7 bits
    - UTF-16
      - 2 o 4 bytes



# Problemas con las codificaciones

---

- ❑ Ejemplo 1:  
Windows con Code::Blocks usando UTF-8
- ❑ Ejemplo 2:  
Windows con Code::Blocks usando la  
codificación predeterminada de Windows  
(CP-1252)
- ❑ Ejemplo 3:  
Linux con Code::Blocks usando UTF-8



```
C:\Users\Latre\Documents\Trabajo\Docencia\prog1f\codigo-c++\practica1\Bienve...
Bienvenidos a la Universidad
Bienvenidos a Programaci||n 1

Process returned 0 (0x0)   execution time : 0.011 s
Press any key to continue.
```

**Símbolo del sistema: CP-850**

**Code::Blocks: UTF-8**

```
5  /*
6  * Pre: ---
7  * Post: Escribe por pantalla el mensaje "Bienvenidos a la Universidad"
8  */
9  int main() {
10     // primera instrucción
11     std::cout << "Bienvenidos a la Universidad" << std::endl;
12     std::cout << "Bienvenidos a Programación 1" << std::endl;
13     // segunda instrucción
14     return 0;
15 }
16
```

UTF-8

# Carácter «ó»

```
circulo.cc  bienvenida.cc x  circunferencia.cc
1  #include <iostream>
2  using namespace std;
3
4  /*
5   * Pre: ---
6   * Post: Escribe por pantalla los mensajes
7   *        "Bienvenidos a la Universidad" y
8   *        "Bienvenidos a Programación 1".
9   */
10 int main() {
11     cout << "Bienvenidos a la Universidad" << endl;
12     cout << "Bienvenidos a Programación 1" << endl;
13     return 0;
14 }
```

C:\ /Bienvenida

Bienvenidos a la Universidad  
Bienvenidos a Programación 1  
Press any key to continue.



# Carácter «ó»

---

- Unicode:

- «ó»

- **Descripción:** Letra latina O minúscula con acento agudo
    - **Punto de código:** U+00F3 (en decimal: 243)
    - **Codificación en UTF-8:** bytes 195 y 179





# Página de códigos 850

|     |   |     |   |     |   |     |   |     |   |     |   |     |   |     |    |
|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|-----|----|
| 128 | Ç | 129 | ü | 130 | é | 131 | â | 132 | ä | 133 | à | 134 | å | 135 | ç  |
| 136 | ê | 137 | ë | 138 | è | 139 | ï | 140 | î | 141 | ì | 142 | Ä | 143 | Å  |
| 144 | É | 145 | æ | 146 | Æ | 147 | ô | 148 | ö | 149 | ò | 150 | û | 151 | ù  |
| 152 | ÿ | 153 | Ö | 154 | Ü | 155 | ø | 156 | £ | 157 | Ø | 158 | × | 159 | f  |
| 160 | á | 161 | í | 162 | ó | 163 | ú | 164 | ñ | 165 | Ñ | 166 | a | 167 | o  |
| 168 | ¿ | 169 | ® | 170 | ¬ | 171 | ½ | 172 | ¼ | 173 | ¡ | 174 | « | 175 | »  |
| 176 | ☼ | 177 | ☼ | 178 | ☼ | 179 |   | 180 | † | 181 | Á | 182 | Â | 183 | Ã  |
| 184 | © | 185 | ‡ | 186 | ‡ | 187 | ‡ | 188 | ‡ | 189 | ç | 190 | ¥ | 191 | ‡  |
| 192 | L | 193 | ⊥ | 194 | ⊥ | 195 | ⊥ | 196 | — | 197 | † | 198 | ã | 199 | Ä  |
| 200 | ℓ | 201 | ℓ | 202 | ℓ | 203 | ℓ | 204 | ℓ | 205 | = | 206 | ‡ | 207 | α  |
| 208 | ð | 209 | Ð | 210 | Ê | 211 | Ë | 212 | È | 213 | ı | 214 | Í | 215 | Î  |
| 216 | Ï | 217 | Ɔ | 218 | Ɔ | 219 | ■ | 220 | ■ | 221 | ı | 222 | Ì | 223 | ■  |
| 224 | Ó | 225 | ß | 226 | Ô | 227 | Ò | 228 | õ | 229 | Õ | 230 | μ | 231 | ρ  |
| 232 | ρ | 233 | Ú | 234 | Û | 235 | Ú | 236 | ý | 237 | Ý | 238 | - | 239 | ˆ  |
| 240 |   | 241 | ± | 242 | — | 243 | ¾ | 244 | ¶ | 245 | § | 246 | ÷ | 247 | ˆ  |
| 248 | ° | 249 | ˆ | 250 | · | 251 | ¹ | 252 | ³ | 253 | ² | 254 | ■ | 255 | 18 |

# Carácter «ó»

```
circulo.cc  bienvenida.cc x  circunferencia.cc
1  #include <iostream>
2  using namespace std;
3
4  /*
5   * Pre: ---
6   * Post: Escribe por pantalla los mensajes
7   *        "Bienvenidos a la Universidad" y
8   *        "Bienvenidos a Programación 1".
9   */
10 int main() {
11     cout << "Bienvenidos a la Universidad" << endl;
12     cout << "Bienvenidos a Programación 1" << endl;
13     return 0;
14 }
```

C:\ /Bienvenida

Bienvenidos a la Universidad  
Bienvenidos a Programación 1  
Press any key to continue.

# Problemas con las codificaciones

---

- Ejemplo 1:  
Windows con Code::Blocks usando UTF-8
- Ejemplo 2:  
Windows con Code::Blocks usando la  
codificación predeterminada de Windows  
(CP-1252)
- Ejemplo 3:  
Linux con Code::Blocks usando UTF-8



```
C:\Users\Latre\Documents\Trabajo\Docencia\prog1f\codigo-c++\practica1\Bienvenida\...  
Bienvenidos a la Universidad  
Bienvenidos a Programaci%  
Process returned 0 (0x0)   execution time : 0.032 s  
Press any key to continue.
```

**Símbolo del sistema: CP-850**

**Code::Blocks: CP-1252**

```
Bienvenida  
Sources  
bienvenida.cpp  
Circunferencia  
Circulo  
Formatear  
3 using namespace std;  
4  
5 /*  
6  * Pre: ---  
7  * Post: Escribe por pantalla el mensaje "Bienvenidos a la Universidad"  
8  */  
9 int main() {  
10     // primera instrucción  
11     std::cout << "Bienvenidos a la Universidad" << std::endl;  
12     std::cout << "Bienvenidos a Programación 1" << std::endl;  
13     // segunda instrucción  
14     return 0;  
15 }  
16  
C/C++ Windows (CR+LF) WINDOWS-1252 line 12, Col 62, Pos 303 Insert Read/Write defa
```

# Carácter «ó»

```
C:\Users\Latre\Documents\Trabajo\Docencia\pr  
Bienvenidos a la Universidad  
Bienvenidos a Programación 1  
Process returned 0 (0x0)   execution  
Press any key to continue.
```

```
circulo.cc  bienvenida.cc x  circunferencia.cc  
1  #include <iostream>  
2  using namespace std;  
3  
4  /*  
5   * Pre: ---  
6   * Post: Escribe por pantalla Los mensajes  
7   *       "Bienvenidos a la Universidad" y  
8   *       "Bienvenidos a Programación 1".  
9   */  
10 int main() {  
11     cout << "Bienvenidos a la Universidad" << endl;  
12     cout << "Bienvenidos a Programación 1" << endl;  
13     return 0;  
14 }
```



# Carácter «ó»

---

- Windows-1252 o CP-1252:
  - «ó»
    - Código entero 243 (00F3 en hexadecimal)





# Página de códigos 850

|     |   |     |   |     |   |     |   |     |   |     |   |     |   |     |    |
|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|-----|----|
| 128 | Ç | 129 | ü | 130 | é | 131 | â | 132 | ä | 133 | à | 134 | å | 135 | ç  |
| 136 | ê | 137 | ë | 138 | è | 139 | ï | 140 | î | 141 | ì | 142 | Ä | 143 | Å  |
| 144 | É | 145 | æ | 146 | Æ | 147 | ô | 148 | ö | 149 | ò | 150 | û | 151 | ù  |
| 152 | ÿ | 153 | Ö | 154 | Ü | 155 | ø | 156 | £ | 157 | Ø | 158 | × | 159 | f  |
| 160 | á | 161 | í | 162 | ó | 163 | ú | 164 | ñ | 165 | Ñ | 166 | a | 167 | o  |
| 168 | ¿ | 169 | ® | 170 | ¬ | 171 | ½ | 172 | ¼ | 173 | ¡ | 174 | « | 175 | »  |
| 176 | ☼ | 177 | ☼ | 178 | ☼ | 179 |   | 180 | ┌ | 181 | Á | 182 | Â | 183 | Ã  |
| 184 | © | 185 | ¶ | 186 |   | 187 | ¶ | 188 | ¶ | 189 | ç | 190 | ¥ | 191 | ¶  |
| 192 | L | 193 | ⊥ | 194 | T | 195 | └ | 196 | — | 197 | ⊕ | 198 | ã | 199 | Ä  |
| 200 | ℓ | 201 | ℓ | 202 | ℓ | 203 | ℓ | 204 | ℓ | 205 | = | 206 | ℓ | 207 | α  |
| 208 | ð | 209 | Ð | 210 | Ê | 211 | Ë | 212 | È | 213 | ı | 214 | Í | 215 | Î  |
| 216 | Ï | 217 | J | 218 | Г | 219 | ■ | 220 | ■ | 221 | ı | 222 | Ì | 223 | ■  |
| 224 | Ó | 225 | ß | 226 | Ô | 227 | Ò | 228 | õ | 229 | Õ | 230 | μ | 231 | ρ  |
| 232 | ρ | 233 | Ú | 234 | Û | 235 | Ù | 236 | ý | 237 | Ý | 238 | - | 239 | ¿  |
| 240 |   | 241 | ± | 242 | — | 243 | ¾ | 244 | ¶ | 245 | § | 246 | ÷ | 247 | ,  |
| 248 | ° | 249 | ¨ | 250 | . | 251 | ¹ | 252 | ³ | 253 | ² | 254 | ■ | 255 | 24 |



# Problemas con las codificaciones

---

- ❑ Ejemplo 1:  
Windows con Code::Blocks usando UTF-8
- ❑ Ejemplo 2:  
Windows con Code::Blocks usando la  
codificación predeterminada de Windows  
(CP-1252)
- ❑ Ejemplo 3:  
Linux con Code::Blocks usando UTF-8



```
Bienvenida

Archivo  Editar  Ver  Terminal  Pestañas  Ayuda

Bienvenidos a Programación 1

Process returned 0 (0x0)   execution time : 0.142 s
Press ENTER to continue.
```

```
bienvenida.cpp [Bienvenida] - Code::Blocks 16.01
File Edit View Search Project Build Debug Tools Plugins Settings Help

<global> main(): int

Management
Projects
  Práctica 1
    Bienvenida
      Sources
        bienvenida.cpp
  Circunferencia
  Circulo
  Formatear

bienvenida.cpp x  circunferencia.cpp x  circulo.cpp x  formatear.cpp x

1  #include <iostream>
2
3  using namespace std;
4
5  /*
6   * Pre: ---
7   * Post: Escribe por pantalla el mensaje "Bienvenidos a Programación 1"
8   */
9  int main() {
10     // primera instrucción
11     std::cout << "Bienvenidos a Programación 1" << std::endl;
12     // segunda instrucción
13     return 0;
14 }
15
```

/home/latre/Documents/Prog1f/codigo-c++/practica1/Bienvenida/bienvenida.cpp Unix (LF) UTF-8 Line 11, Column 47 Insert Read/Wri... default

# Más información

---

- Joel Spolsky, «The Absolute Minimum Every Software Developer Absolutely, Positively Must Know About Unicode and Character Sets (No Excuses!)», *Joel on Software*, 8-10-2013.
- <https://www.joelonsoftware.com/2003/10/08/the-absolute-minimum-every-software-developer-absolutely-positively-must-know-about-unicode-and-character-sets-no-excuses/>



# Operadores asociados

---

- Los de los tipos enteros
  - Aritméticos: +, -, ...
  - Relación: ==, !=, <, <=, >, >=
- Conversión con enteros pueden ser explícitas:
  - `int( 'A' )` devuelve 65
  - `char(66)` devuelve 'B'
- Hay secuencias de caracteres con códigos consecutivos crecientes:
  - Mayúsculas del alfabeto inglés: 'A', 'B', 'C', ..., 'X', 'Y' y 'Z'
  - Minúsculas del alfabeto inglés: 'a', 'b', 'c', ..., 'x', 'y' y 'z'
  - Dígitos: '0', '1', '2', '3', '4', '5', '6', '7', '8' y '9'

# Expresiones con caracteres

---

- `char c = 'E';`
- `c == 'A'`
- `c != 'e'`
- `c >= 'A'`
- `c <= 'Z'`
- `c >= 'A' && c <= 'Z'`
- `c >= 'a'`
- `c <= 'z'`
- `c >= 'a' && c <= 'z'`
- `char(c + 1)`
- `char(c + 32)`
- `char(c - 'A' + 'a')`

# Ejemplos

```
/*  
 * Pre: ---  
 * Post: Si «c» es un carácter que representa  
 * una letra en mayúscula entonces  
 * devuelve true; en otro caso devuelve  
 * false.  
 */  
bool esMayuscula(const char c) {  
    return c >= 'A' && c <= 'Z';  
}
```

# Ejemplos

```
/*  
 * Pre: ---  
 * Post: Si «c» es un carácter que  
 * representa una letra minúscula  
 * entonces devuelve true; en otro  
 * caso devuelve false.  
 */  
bool esMinuscula(const char c) {  
    return c >= 'a' && c <= 'z';  
}
```

# Ejemplos

```
/*  
 * Pre: ---  
 * Post: Si «c» es un carácter que  
 * representa un dígito entonces  
 * devuelve true; en otro caso  
 * devuelve false.  
 */  
bool esDigito(const char c) {  
    return c >= '0' && c <= '9';  
}
```



# Biblioteca estándar ctype

---

- Character handling functions. This header declares a set of functions to classify and transform individual characters.
  - `isalnum`: Check if character is alphanumeric
  - `isalpha`: Check if character is alphabetic
  - `isblank`: Check if character is blank
  - `isdigit`: Check if character is decimal digit
  - `islower`: Check if character is lowercase letter
  - `ispunct`: Check if character is a punctuation character
  - `isspace`: Check if character is a white-space
  - `isupper`: Check if character is uppercase letter
  - `tolower`: Convert uppercase letter to lowercase
  - `toupper`: Convert lowercase letter to uppercase

# Ejemplos

```
/*  
 * Pre:  ---  
 * Post: Si «c» es un carácter que representa un dígito entonces  
 *        devuelve el valor numérico comprendido entre 0 y 9  
 *        representado por «c»;  
 *        en otro caso devuelve un valor negativo.  
 */  
int valorDigito(const char c) {  
    if (esDigito(c)) {  
        return c - '0';  
    }  
    else {  
        return -1;  
    }  
}
```

# Cadenas de caracteres

---

- ❑ Secuencia de 0, 1 o más caracteres
- ❑ Representación literal entre comillas
  - ""
  - "A"
  - "Programación 1"
- ❑ Tipos de datos para su representación
  - Tabla de datos de tipo **char**
  - Clase predefinida `string`
    - ❑ No la usaremos en este curso

# Cadenas de caracteres

```
char nombre[] = { 'M', 'a', 'n', 'u', 'e', 'l', '\0' };  
char apellidos[] = "Rodrigo Merino";
```

nombre

|   |   |   |   |   |   |     |
|---|---|---|---|---|---|-----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6   |
| M | a | n | u | e | l | NUL |

apellidos

|   |   |   |   |   |   |   |   |   |   |    |    |    |    |     |
|---|---|---|---|---|---|---|---|---|---|----|----|----|----|-----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14  |
| R | o | d | r | i | g | o |   | M | e | r  | i  | n  | o  | NUL |

nombre

|   |     |
|---|-----|
| 0 | M   |
| 1 | a   |
| 2 | n   |
| 3 | u   |
| 4 | e   |
| 5 | l   |
| 6 | NUL |

apellidos

|    |     |
|----|-----|
| 0  | R   |
| 1  | o   |
| 2  | d   |
| 3  | r   |
| 4  | i   |
| 5  | g   |
| 6  | o   |
| 7  |     |
| 8  | M   |
| 9  | e   |
| 10 | r   |
| 11 | i   |
| 12 | n   |
| 13 | o   |
| 14 | NUL |

# Cadenas de caracteres

```
#include <iostream>
using namespace std;
const int MAX_LONG_NOMBRE = 20;

/*
 * Pre:  ---
 * Post: He pedido el nombre del usuario, lo ha leído del
 *       teclado y lo ha vuelto a escribir en la pantalla.
 */
int main() {
    char nombre[MAX_LONG_NOMBRE];
    cout << "Escriba su nombre: " << flush;
    cin >> nombre;
    cout << "Su nombre es: " << nombre << endl;
    return 0;
}
```

# Cadenas de caracteres

```
#include <iostream>
using namespace std;
const int MAX_LONG_NOMBRE = 20;

/*
 * Pre: ---
 * Post: He pedido el nombre al usuario
 *       teclado
 */

int main() {
    char nombre[MAX_LONG_NOMBRE];
    cout << "Escriba su nombre: " << flush;
    cin >> nombre;
    cout << "Su nombre es: " << nombre << endl;
    return 0;
}
```

nombre

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|
| ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ?  | ?  | ?  | ?  | ?  | ?  | ?  | ?  | ?  | ?  |

# Cadenas de caracteres

```
#include <iostream>
using namespace std;
const int MAX_LONG_NOMBRE = 20;

/*
 * Pre: ---
 * Post: He pedido el nombre al usuario
 * teclado
 */

int main() {
    char nombre[MAX_LONG_NOMBRE];
    cout << "Escriba su nombre: " << flush;
    cin >> nombre;
    cout << "Su nombre es: " << nombre << endl;
    return 0;
}
```

| nombre |   |   |   |   |   |     |   |   |   |    |    |    |    |    |    |    |    |    |    |
|--------|---|---|---|---|---|-----|---|---|---|----|----|----|----|----|----|----|----|----|----|
| 0      | 1 | 2 | 3 | 4 | 5 | 6   | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| M      | a | n | u | e | l | NUL | ? | ? | ? | ?  | ?  | ?  | ?  | ?  | ?  | ?  | ?  | ?  | ?  |

# Cadenas de caracteres

---

- En la biblioteca predefinida `<cstring>`
  - `strcpy(cad1, cad2)`
    - copia la cadena `cad2` en `cad1`
  - `strcat(cad1, cad2)`
    - modifica `cad1` concatenándole la cadena `cad2`
  - `strlen(cad)`
    - devuelve la longitud de la cadena `cad`
  - `strcmp(cad1, cad2)`
    - compara las cadenas `cad1` y `cad2`
    - devuelve 0 si el contenido de ambas es idéntico
    - devuelve un valor positivo si `cad1` es alfabéticamente posterior a `cad2`
    - devuelve un valor negativo si `cad1` es alfabéticamente anterior a `cad2`



# Ejemplo

```
char letras[] = "ABCD";
char alfabeto[30] = "";
cout << "[" << alfabeto << "]" tiene " << strlen(alfabeto)
      << " caracteres" << endl;

strcpy(alfabeto, letras);
cout << "[" << alfabeto << "]" tiene " << strlen( alfabeto )
      << " caracteres" << endl;

strcat(alfabeto, "EFGHIJ");
cout << "[" << alfabeto << "]" tiene " << strlen( alfabeto )
      << " caracteres" << endl;

strcat(alfabeto, "KLMNOPQRSTUVWXYZ");
cout << "[" << alfabeto << "]" tiene " << strlen( alfabeto )
      << " caracteres" << endl;
```

# Ejemplo

[ ] tiene 0 caracteres

[ABCD] tiene 4 caracteres

[ABCDEFGHIIJ] tiene 10 caracteres

[ABCDEFGHIJKLMNOPQRSTUVWXYZ] tiene 26 caracteres

# Conversión de cadena a entero

```
/*  
 * Pre: «cadena» almacena una secuencia de caracteres  
 * que representan literalmente un entero con la  
 * siguiente sintaxis:  
 * <literal_entero> ::= [ <signo> ]  
 *                         <digito> { <digito> }  
 * <signo> ::= "+" | "-"  
 * <digito> ::= "0" | "1" | "2" | "3" | "4"  
 *               | "5" | "6" | "7" | "8" | "9"  
 * Post: Devuelve el valor entero representado en  
 * «cadena».  
 */  
int valorEntero(const char cadena[]);
```

# Conversión de cadena a entero

```
int valorEntero(const char cadena[]) {  
    int i = 0;  
    bool negativo = false;  
  
    // cadena[0] es un dígito,  
    // un signo más o un signo menos  
    if (cadena[0] == '+') {  
        i++;  
    }  
    else if (cadena[0] == '-') {  
        i++;  
        negativo = true;  
    }  
  
    ...  
}
```

# Conversión de cadena a entero

```
int valorEntero(const char cadena[]) {  
    // cadena[i] es un dígito  
    int valor = 0;  
    while (cadena[i] != '\0') {  
        int valorDigito = cadena[i] - '0';  
        valor = 10 * valor + valorDigito;  
        i++;  
    }  
    if (negativo) {  
        valor = -valor;  
    }  
    return valor;  
}
```

# Conversión de entero a cadena

```
/*
 * Pre: ---
 * Post: Asigna a «literal» una secuencia de caracteres
 * que representa literalmente el entero «valor»
 * con la siguiente sintaxis:
 *     <literal_entero> ::= [ <signo> ]
 *                          <digito> { <digito> }
 *     <signo> ::= "+" | "-"
 *     <digito> ::= "0" | "1" | "2" | "3" | "4"
 *                | "5" | "6" | "7" | "8" | "9"
 */
void literalEntero(const int valor, char literal[]);
```

# Conversión de entero a cadena (versión apuntes)

```
void literalEntero(const int valor, char literal[]) {  
    // Determinación de si valor es negativo o no  
  
    // Cálculo de las unidades. Fuera del bucle para  
    // que la representación de 0 sea "0".  
  
    // Cálculo del resto de las cifras de menos  
    // a más significativas  
  
    // Se pone el signo si es preciso  
  
    // Inversión de los dígitos almacenados en  
    // «literal»  
  
    // Inserta el carácter de final de la cadena  
}
```



# Conversión de entero a cadena (versión apuntes)

```
void literalEntero(const int valor, char literal[]) {  
    // Determinación de si valor es negativo o no  
    bool negativo = false;  
    if (valor < 0) {  
        negativo = true;  
        valor = -valor;  
    }  
  
    // Cálculo de las unidades. Fuera del bucle  
    // para que la representación de 0 sea "0".  
    literal[0] = '0' + valor % 10;  
    valor = valor / 10;  
  
    ...  
}
```



# Conversión de entero a cadena (versión apuntes)

```
void literalEntero(const int valor, char literal[]) {  
    ...  
    // Cálculo del resto de las cifras de menos a más  
    // significativas. El cursor «i» indica la siguiente  
    // componente de «literal» a almacenar.  
    int i = 1;  
    while (valor != 0) {  
        literal[i] = '0' + valor % 10;  
        valor = valor / 10;  
        i++;  
    }  
  
    // Se pone el signo si es preciso  
    if (negativo) {  
        literal[i] = '-';  
        i++;  
    }  
    ...  
}
```

# Conversión de entero a cadena (versión apuntes)

```
void literalEntero(const int valor, char literal[])  
{  
    ...  
  
    // Inversión de los dígitos almacenados  
    // en «literal»  
    for (int j = 0; j < i / 2; j++) {  
        char aux = literal[j];  
        literal[j] = literal[i - j - 1];  
        literal[i - j - 1] = aux;  
    }  
  
    // Inserta el carácter de final de la cadena  
    literal[i] = '\\0';  
}
```

# Conversión de entero a cadena (otra versión)

```
void literalEntero(const int valor,  
                  char literal[]) {  
    // Cálculo del número de cifras.  
    // Determinación de si valor es negativo o no.  
    // Inserción del carácter de final de la  
    // cadena.  
    // Cálculo de las cifras de menos a más  
    // significativas, almacenándose en las  
    // componentes finales.  
}
```

# Conversión de entero a cadena (otra versión)

```
void literalEntero(const int valor, char literal[]) {  
    // Cálculo del número de cifras  
    int i = numCifras(valor);  
    int n = valor;  
  
    // Determinación de si valor es negativo o no  
    if (valor < 0) {  
        literal[0] = '-';  
        n = -valor;  
        i++;  
    }  
    // Inserta el carácter de final de la cadena  
    literal[i] = '\\0';  
    ...  
}
```

# Conversión de entero a cadena (otra versión)

```
void literalEntero(const int valor, char literal[]) {  
    ...  
  
    if (n == 0) {  
        // Tratamiento específico del caso del 0;  
        literal[0] = '0';  
    }  
    else {  
        // Cálculo de las cifras de «n» de menos a más  
        // significativas  
        while (n > 0) {  
            i--;  
            literal[i] = '0' + n % 10;  
            n = n / 10;  
        }  
    }  
}
```