

Tabla ASCII

Karl Heitmann

Academia Desafío Latam

2020

Introducción

- ▶ Los computadores son unas máquinas para procesar información.

Introducción

- ▶ Los computadores son unas máquinas para procesar información.
- ▶ Garbage in \longrightarrow Garbage out

Introducción

- ▶ Los computadores son unas máquinas para procesar información.
- ▶ Garbage in \longrightarrow Garbage out
- ▶ La Tabla ASCII es un estándar para representar **información**.

Introducción

- ▶ Los computadores son unas máquinas para procesar información.
- ▶ Garbage in \longrightarrow Garbage out
- ▶ La Tabla ASCII es un estándar para representar **información**.
- ▶ Los computadores entienden un lenguaje especial. Un código especial.

Introducción

- ▶ Los computadores son unas máquinas para procesar información.
- ▶ Garbage in → Garbage out
- ▶ La Tabla ASCII es un estándar para representar **información**.
- ▶ Los computadores entienden un lenguaje especial. Un código especial.
- ▶ Pero antes... ¡vamos a ver un poco de historia!

Introducción - Matemáticas

- ▶ Nosotros usamos el sistema decimal y los número arábigos para las matemáticas: del 0 al 9.

Introducción - Matemáticas

- ▶ Nosotros usamos el sistema decimal y los número arábigos para las matemáticas: del 0 al 9.



Figure: tenemos un sistema decimal porque contamos con los dedos

- ▶ Manfred Spitzer Demencia Digital — Chinos contando con las manos

Introducción - Giuseppe Peano

Giuseppe Peano (1858 - 1932),
matemático italiano, estudió los
números naturales.



- ▶ Él los define como una secuencia, como el número que sigue al siguiente.

Introducción - Giuseppe Peano

Giuseppe Peano (1858 - 1932), matemático italiano, estudió los números naturales.



- ▶ Él los define como una secuencia, como el número que sigue al siguiente.
- ▶ Son conocidos los axiomas de Peano, que se han usado en investigaciones de metamatemáticas, los fundamentos de la matemática.

Código binario

- ▶ Solamente existen dos figuras para representar los números: el uno y el cero: 1 - 0

Código binario

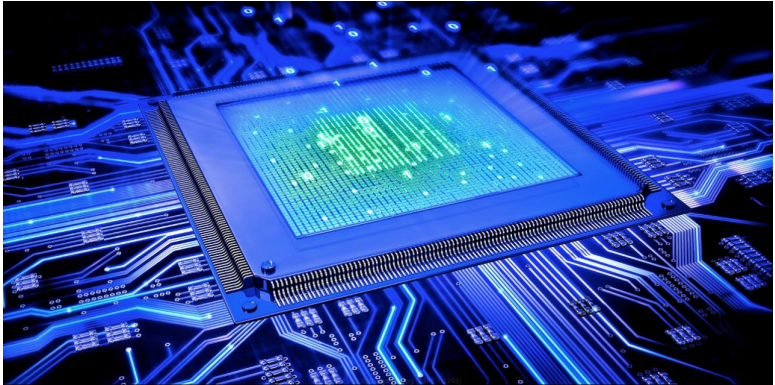
- ▶ Solamente existen dos figuras para representar los números: el uno y el cero: 1 - 0
- ▶ El cero representa ausencia de electricidad, y el uno representa presencia de electricidad.

Código binario

- ▶ Solamente existen dos figuras para representar los números: el uno y el cero: 1 - 0
- ▶ El cero representa ausencia de electricidad, y el uno representa presencia de electricidad.
- ▶ Los electrones



CPU



Como es una instrucción binaria

MSR (transfer register contents to PSR)

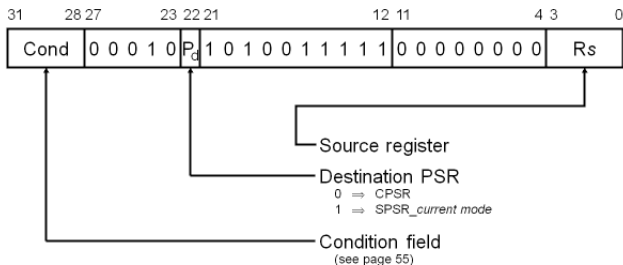


Figure: Sacado de RISC OS

¿Y qué tiene que ver todo esto con la tabla ASCII?

- ▶ CPU donde llega información e instrucciones

¿Y qué tiene que ver todo esto con la tabla ASCII?

- ▶ CPU donde llega información e instrucciones
- ▶ Un catálogo inmenso de instrucciones distintas

¿Y qué tiene que ver todo esto con la tabla ASCII?

- ▶ CPU donde llega información e instrucciones
- ▶ Un catálogo inmenso de instrucciones distintas
- ▶ Los resultados se deben mostrar al usuario (y los pasos intermedios también, en caso de *debuguear* el código)

¿Y qué tiene que ver todo esto con la tabla ASCII?

- ▶ CPU donde llega información e instrucciones
- ▶ Un catálogo inmenso de instrucciones distintas
- ▶ Los resultados se deben mostrar al usuario (y los pasos intermedios también, en caso de *debuguear* el código)
- ▶ Aquí es donde llega la tabla ASCII

American Standard Code for Information Interchange

- ▶ Aparece en 1963, considerada una milestone o un hito en la historia de la electrónica según la IEEE

American Standard Code for Information Interchange

- ▶ Aparece en 1963, considerada una milestone o un hito en la historia de la electrónica según la IEEE
- ▶ Elaborado por la ANSI: ~~American Standard Code for Information Interchange.~~

American National Standards Institute

American Standard Code for Information Interchange

- ▶ Aparece en 1963, considerada una milestone o un hito en la historia de la electrónica según la IEEE
- ▶ Elaborado por la ANSI: ~~American Standard Code for Information Interchange~~.
- ▶ Tardó 3 años en formarse el estándar
American National Standards Institute

American Standard Code for Information Interchange

- ▶ Aparece en 1963, considerada una milestone o un hito en la historia de la electrónica según la IEEE
- ▶ Elaborado por la ANSI: ~~American Standard Code for Information Interchange~~.
- ▶ Tardó 3 años en formarse ~~American~~ American National Standards Institute
- ▶ Usa 7 bits para representar los caracteres. Esto da 128 caracteres distintos.

American Standard Code for Information Interchange

- ▶ Aparece en 1963, considerada una milestone o un hito en la historia de la electrónica según la IEEE
- ▶ Elaborado por la ANSI: ~~American Standard Code for Information Interchange~~.
- ▶ Tardó 3 años en formarse. ~~American~~ National Standards Institute
- ▶ Usa 7 bits para representar los caracteres. Esto da 128 caracteres distintos.

!"#\$%&'()*+,-./012
3456789:;<=>?@ABCDE
FGHIJKLMNOPQRSTUVWXYZ
[\]^_`abcdefghijklmnopqrstuvwxyz
{|}~

Figure: Lista de caracteres imprimibles ASCII. 95 caracteres

Orden de los caracteres

- ▶ Se compone de 95 caracteres imprimibles y de 33 códigos de control (el "line feed", el "carriage return" y el "tab" se siguen utilizando hoy en día, el resto está obsoleto)

Orden de los caracteres

- ▶ Se compone de 95 caracteres imprimibles y de 33 códigos de control (el "line feed", el "carriage return" y el "tab" se siguen utilizando hoy en día, el resto está obsoleto)
- ▶ Mayúsculas vienen antes que las minúsculas

Orden de los caracteres

- ▶ Se compone de 95 caracteres imprimibles y de 33 códigos de control (el "line feed", el "carriage return" y el "tab" se siguen utilizando hoy en día, el resto está obsoleto)
- ▶ Mayúsculas vienen antes que las minúsculas
- ▶ Signos de puntuación y dígitos vienen antes de las letras.

Características notables

- ▶ Permite transformar de mayúsculas a minúsculas y vice versa con una simple adición.

Características notables

- ▶ Permite transformar de mayúsculas a minúsculas y vice versa con una simple adición.
- ▶ El orden en la tabla permite ordenar frases alfabéticamente.

Características notables

- ▶ Permite transformar de mayúsculas a minúsculas y vice versa con una simple adición.
- ▶ El orden en la tabla permite ordenar frases alfabéticamente.
- ▶ Requiere ingenio elaborar algo así, y está apoyado sobre las matemáticas.

Pero tiene algunos problemas...

- ▶ Está basado en el idioma inglés. Es excelente para el inglés, pero no es muy bueno para el francés o el castellano.

Pero tiene algunos problemas...

- ▶ Está basado en el idioma inglés. Es excelente para el inglés, pero no es muy bueno para el francés o el castellano.
- ▶ Aparecen nuevas configuraciones (encoding) de caracteres que extienden el alcance de la tabla ASCII. Los más populares son el ISO8859 (latin) y unicode (UTF-8).

ISO/IEC - ISO8859

- ▶ ISO: *International Organization for Standardization*

ISO/IEC - ISO8859

- ▶ ISO: *International Organization for Standardization*
- ▶ IEC: *International Electrotechnical Commission*

ISO/IEC - ISO8859

- ▶ ISO: *Internation Organization for Standardization*
- ▶ IEC: *International Electrotechnical Commission*
- ▶ Diseñado para un intercambio confiable de información.
Ocupa 8 bits de información, en lugar de los 7 bits de información que ocupa la tabla ASCII

ISO/IEC - ISO8859

- ▶ ISO: *International Organization for Standardization*
- ▶ IEC: *International Electrotechnical Commission*
- ▶ Diseñado para un intercambio confiable de información. Ocupa 8 bits de información, en lugar de los 7 bits de información que ocupa la tabla ASCII
- ▶ Se compone de 16 partes que se han ido agregando desde el año 1987 al año 2001.

ISO/IEC - ISO8859

- ▶ ISO: *International Organization for Standardization*
- ▶ IEC: *International Electrotechnical Commission*
- ▶ Diseñado para un intercambio confiable de información. Ocupa 8 bits de información, en lugar de los 7 bits de información que ocupa la tabla ASCII
- ▶ Se compone de 16 partes que se han ido agregando desde el año 1987 al año 2001.
- ▶ En cada parte nueva se cambian algunos caracteres por otros, para dar soporte a lenguas específicas.

Unicode

- ▶ Cambia el modelo que se usaba hasta el estándar ISO8859: utiliza un protocolo.

Unicode

- ▶ Cambia el modelo que se usaba hasta el estándar ISO8859: utiliza un protocolo.
- ▶ Utiliza una secuencia de bytes para codificar los caracteres.

Unicode

- ▶ Cambia el modelo que se usaba hasta el estándar ISO8859: utiliza un protocolo.
- ▶ Utiliza una secuencia de bytes para codificar los caracteres.
- ▶ Existen UTF-8 UTF-16 y UTF-32

Unicode

- ▶ Cambia el modelo que se usaba hasta el estándar ISO8859: utiliza un protocolo.
- ▶ Utiliza una secuencia de bytes para codificar los caracteres.
- ▶ Existen UTF-8 UTF-16 y UTF-32
- ▶ Lo que busca es codificar todos los caracteres del mundo: incluso ha permitido codificar los jeroglifos egipcios

Unicode

- ▶ Cambia el modelo que se usaba hasta el estándar ISO8859: utiliza un protocolo.
- ▶ Utiliza una secuencia de bytes para codificar los caracteres.
- ▶ Existen UTF-8 UTF-16 y UTF-32
- ▶ Lo que busca es codificar todos los caracteres del mundo: incluso ha permitido codificar los jeroglifos egipcios
- ▶ Ha permitido hasta incluir emojis en los textos.

Unicode

- ▶ Cambia el modelo que se usaba hasta el estándar ISO8859: utiliza un protocolo.
- ▶ Utiliza una secuencia de bytes para codificar los caracteres.
- ▶ Existen UTF-8 UTF-16 y UTF-32
- ▶ Lo que busca es codificar todos los caracteres del mundo: incluso ha permitido codificar los jeroglifos egipcios
- ▶ Ha permitido hasta incluir emojis en los textos.
- ▶ Hay Kata en Codewars donde se plantea el desafío de implementar un encoder y un decoder unicode...

# BYTES	FIRST CODE POINT	LAST CODE POINT	BYTE 1	BYTE 2	BYTE 3	BYTE 4
1	0	127	0xxxxxxx			
2	128	2047	110xxxxx	10xxxxxx		
3	2048	65535	1110xxxx	10xxxxxx	10xxxxxx	
4	65536	1114111	11110xxx	10xxxxxx	10xxxxxx	10xxxxxx

Cierre

- ▶ Las matemáticas son el fundamento.

Cierre

- ▶ Las matemáticas son el fundamento.
- ▶ Los caracteres que vemos son el resultado de una evolución. Aparecen ideas, las buenas se popularizan y se adoptan.

Cierre

- ▶ Las matemáticas son el fundamento.
- ▶ Los caracteres que vemos son el resultado de una evolución. Aparecen ideas, las buenas se popularizan y se adoptan.
- ▶ Unicode, ¿qué vendrá después?