

NAME	PAGES	SPEAKER/CLASS	DATE - TIME
Miguel Angel Cárdenas	1	Viernes	3/11/25

Title: Sistema de Codificación de Caracteres (ASCII)

Keyword	Topic: ASCII
interchange Caracteres ASCII	<p>Notes: Historia: El Código ASCII (símbolos en máquinas) American Standard Code for information interchange, es decir Código Americano.</p> <p>Fue creado en 1963 por el Comité estadounidense de estándares o (ASA), este organismo cambió su nombre en 1969 por "Instituto estadounidense de estándares nacionales" o (ANSI).</p>
Questions	<p>Este Código nació a partir de reordenar y expandir el Conjunto de Símbolos y caracteres ya utilizados en aquél momento en telégrafos por la Compañía Bell. En un primer momento solo incluía Letras mayúsculas y números, pero en 1967 se incorporaron las letras minúsculas y algunos caracteres de control formando así lo que se conoce como US-ASCII, es decir los caracteres del 0 al 127.</p>

Summary: El Código ASCII reserva los primeros 32 códigos, numerados del 0 al 31 en decimal) para caracteres de control (códigos no pensados originalmente para representar información imprimible).

NAME	PAGES	SPEAKER/CLASS	DATE - TIME
miguel angel Caceres	2	viernes	31/11/25

Title: Sistema de Codificación de Caracteres (ASCII)

Keyword

Ambiguedad
bits
codificar

Topic: Cont.

Notes: ASCII utiliza 7 bits para representar $2^7 = 128$ caracteres

			Binario
A	L. mayúscula A	Código 65	0100001
Z	L. mayúscula Z	Código 90	01011010
a	L. minúscula a	Código 97	0110001
z	L. minúscula z	Código 122	01111010
0	Dígito 0	Código 48	00110000
9	Dígito 9	Código 57	00111001

Questions

Ejemplo: para codificar la palabra HOLA en ASCII, se toma el código decimal o binario de cada carácter individual.

Carácter	Decimal	Binario
H	72	01001000
O	79	01101111
L	108	01101100
A	97	01100001

por lo tanto, la palabra "HOLA" se representa en binario ASCII como lo

Summary: Recencia de bytes:

01001000 01101111 01101100 01100001

El sistema ASCII establece un estandar universal para que los computadores puedan intercambiar texto. Sin Ambigüedad

NAME	PAGES	SPEAKER/CLASS	DATE/TIME
miguel angel colares	3	viernes	3/11/25

Title: Sistema de codificación de caracteres UTF-8

Keyword

unicode
retrocompatibilidad
caracter

Topic: UTF-8

Notes: (UTF-8) es un sistema de codificación de caracteres de ancho variable que representa cualquier carácter del estándar unicode. Su principal ventaja es la retrocompatibilidad con ASCII, ya que los primeros 128 caracteres son idénticos y ocupan un solo byte.

Esto hace que UTF-8 sea eficiente para texto en inglés y otros alfabetos latinos latinos, mientras que caracteres más complejos, como árabes y otros idiomas, pueden ocupar de 2 a 4 bytes.

Características principales de UTF-8

- Ancho variable: Admite un número de bytes entre 1 y 4 para representar cada punto de código (carácter) unicode.
- Eficiencia: Utiliza menos byte para caracteres más frecuentes, como los del

Summary: inglés, lo que ahorra espacio de almacenamiento.

- Auto-sincronizada: permite buscar y decodificar un carácter desde cualquier punto de la secuencia de datos sin necesidad de procesar toda la cadena.

By Carlos Pichardo Vilque

NAME	PAGES	SPEAKER/CLASS	DATE - TIME
Miguel Angel Cárceles	2	viernes	3/11/25

Title: Sistemas de codificación de carácter (UTF-8)

Keyword

Topic: Cont...

Notes: ¿Por qué es tan utilizado?

Es muy usado porque permite la comunicación sin problemas entre personas que usan diferentes idiomas, ya que puede manejar caracteres prácticamente todos los alfabetos.

Ejemplos: Caracteres de 1 Byte (compatibles con ASCII)

Questions

¿Por qué es tan utilizado?

Carácter	Descripción	Decimal	UTF-8
A	Línea A	65	0100 001
!	S. exclamación	33	0010 0001

Caracteres de 2 Bytes (alfabetos extendidos)

Para representar caracteres no presentes en ASCII, como los acentos o el 'ñ' UTF-8 utiliza dos bytes.

Carácter	Descripción	Decimal	UTF-8
ñ	Ent. minis	241	1011 0001

Summary:

Á	A con acento	225	1100 0011
			1010 0001

Caracteres de 3 Bytes (caracteres árabes)

para la mayoría de los caracteres (CJK (chino, japonés, coreano)).

NAME	PAGES	SPEAKER/CLASS	DATE - TIME
miguel angel coceres	3	viernes	3/11/25

Title: sistemas de codificación de caracteres UTF-8

Keyword

Topic: Cont...

Notes: Caracter descripción Decimal UTF-8

JK	Caracter chino	11100110
	poco uso	10110000
		10110100

Caracter de 4 Bytes (emojis)

Los emojis no son caracteres normales comunes
usan cuatro bytes

Caracter Descripción decimal UTF-8

†	Brida	10024	11100000
			10011111
			101001000
			101000000

Questions

Summary:

NAME	PAGES	SPEAKER/CLASS	DATE/TIME
Miguel Ángel Callejas	1	Viernes	3/11/25

Title: Sistema de Codificación de caracteres UTF-16

Keyword

Codificación
unicode

Topic: UTF-16

Notes: Es un sistema de codificación de caracteres que utiliza una o dos unidades de código de 16 bits para representar un carácter. Permite el código unicode. Esto significa que cada carácter ocupa entre 2 y 4 bytes, dependiendo su valor. Para caracteres de uno mismo, como la letra "A" se utiliza una única unidad de 16 bits (2 bytes).

Questions

Características:

- unidad básica de 16 bits: El nombre UTF-16 proviene de la unidad de código de 16 bits (2 bytes).
- longitud variable: La codificación tiene una longitud variable, ya que algunos caracteres necesitan 2 bytes y otros 4 bytes.

Summary: Ventajas: Es una codificación eficiente para texto que utiliza caracteres asciíacos y otros caracteres de uso no extendido, ya que esto al mundo requiere menos espacio que en UTF-8.

NAME	PAGES	SPEAKER/CLASS	DATE - TIME
Miguel Angel Caceres	2	viernes	3/11/25

Title: Sistemas de codificación de caracteres UTF-16

Keyword

Subrayado

Topic: Cont...

Notes: Desventaja: puede ser menos eficiente en terminos de espacio que UTF-8 para texto que usa principalmente caracteres ASCII, ya que todos los caracteres ocupan al menos 2 bytes en la medida de 1.

Example: Caracter Basico (2 Bytes)

Questions

Carac. nombre Código (unificado) UTF-16
 (Hexadecimal)
 é Lé 0T00E9 00E9

Caracter extendido (4Bytes)

Para los caracteres fuera del BMP, UTF-16 utiliza un par de subrayados, sumando un total de 4 bytes (32 bits)

Caracter nombre Código (unificado) UTF-16
 ábola 0T1F680 D830 DE80

Summary: