

# Base de Datos

DC. FCEyN

2023-10-25 teórica 11 (segunda parte)

en la primera parte se vio el apunte de logging multiversión

Emilio Platzer  
[tute@dc.uba.ar](mailto:tute@dc.uba.ar)

$\exists \nrightarrow \Rightarrow \forall$

# Textos Caracteres

Codepage -> la lista de caracteres posibles y su código numérico.

- ISO8859-1, ANSI Windows-1251 (256 entradas)

- Unicode (149.000 entradas en la versión 15 y agregando...)

! " # \$ % & ' ( ) \* + , - . / 0 1 2 3 4 5 6 7 8 9 : ; < = > ?  
@ 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 [ \ ] ^ \_  
` 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 { | } ~  
.  
ı ç ₺ ₪ ¥ ¦ § ¨ © ª « ¬ ® ¯ ° ± ² ³ ´ µ ¶ · ¸ ¹ º » ¼ ½ ¾ ¿  
À Á Â Ã Ä Å Æ Ç È É Ê Ë Ì Í Î Ï Ð Ñ Ò Ó Ô Õ Ö × Ø Ù Ú Û Ü Ý Þ ß  
à á â ã ä å æ ç è é ê ë ì í î ï ð ñ ò ó ô õ ö ÷ ø ù ú û ü ý þ ÿ

á à ä ã  
□ α á χ α ς f A b c d e f g h i j k l m n o p r s t u v w  
x y z 1 2 3 4 5 6 7 8 9 0 ♔ ♚ ♜ ♞ ♟ ♠ ♡ ♢ ♣  
東 ☁ ☺ ☺ ☹ ☹

# Text - encoding

- Ansi: 1 letra 1 byte
- Unicode:
  - UTF8 (a veces 1, 2, 3 ó 4 bytes)
  - UTF16 (a veces 2 a veces 4)
  - UTF32 (4 bytes)
  - Especiales
    - BOM
    - subrogados
    - de usuario

# Textos - collation

- Ordenamiento (comparar por <)
  - caracteres numéricos
  - espacio de fin de palabras
  - mayúsculas, minúsculas
  - acentos
  - signos de puntuación
- Identidad (comparar por =)
  - Afecta las PK
  - Afecta los Where
- Idioma
  - ¿La ñ es una n con acento? ¿y la Ç?

# Fechas - problemas

- Existen varios calendarios para fechas pasadas: Gregoriano, Juliano
  - Adopción gradual del calendario gregoriano (varios siglos de superposición)
    - Calendario gregoriano proléptico
- Por las zonas horarias la fecha cambia en distintos momentos
- Hay lenguajes (Java, C#, Javascript) que no tienen un tipo date native
  - Almacenan los Dates en DateTime (y son sensibles a las zonas horarias)
- Las bases de datos SQL sí tienen un tipo date
- Los formatos de texto para fechas se estandarizan en ISO-8601

# Horas - problemas

- Datetimes hay en los lenguajes de programación y en SQL (llamado también timestamp)
- Las zonas horarias se pueden denotar con nombres (para que se actualicen automáticamente cuando cambia el horario de invierno/verano)
- Hay timestamp con y sin marca de zona horaria
  - Hay que elegir bien cuál según el uso
- Los servidores deberían estar sincronizados lo mayor possible
- Cuando cambia la hora hay días de 23hs o de 25hs, no todos los lenguajes lo tratan de la misma manera.
- Cada tanto se agrega un lap second.

# Timestamps en la base de datos

- En general hay 3 funciones que devuelven el timestamp.
- En Postgres:
  - `current_timestamp`: el momento del inicio de la transacción (también `transaction_timestamp`)
  - `sentence_timestamp`: el momento del inicio de la instrucción (del insert, update, delete)
  - `clock_timestamp`: la hora actual

# Medir el tiempo

- Restar dos fechas (o datetimes) devuelve una cantidad de tiempo, no una fecha (ni un datetime). Devuelve una cantidad de tiempo.
- En Postgres (SQL standard) se usa el tipo **interval** que puede almacenar tiempos en microsegundos, segundos, minutos, horas y días en forma precisa. Puede almacenar meses, años en la forma esperada.
- En SQL Server existe el tipo time que puede usarse para almacenar tiempos menores a un día
- ISO-8601 estandariza el formato texto de cantidades de tiempo.
  - Ejemplo “P1Y2M10DT2H30M” se sufijan las partes.