

Tabela ASCII e Unicode

Os caracteres possuem um número inteiro correspondente na tabela ASCII (inglês *American Standard Code for Information Interchange*; português - Código Padrão Americano para o Intercâmbio de Informação). A codificação ASCII é usada para representar textos em computadores, equipamentos de comunicação, entre outros dispositivos que trabalham com texto. Desenvolvida a partir de 1960, grande parte das codificações de caracteres modernas a herdaram como base. A partir da tabela ASCII outros tipos de tabelas foram criadas, por exemplo, a Unicode que é um superconjunto ASCII. Os números existentes na tabela ASCII tem o mesmo significado na tabela Unicode. Por exemplo, o número 65 significa 'A' em ambas tabelas.

Você não pode salvar texto em seu disco rígido como "Unicode", porque é uma representação abstrata do texto. Você precisa "codificar" esta representação abstrata, através da codificação UTF-8, UTF-16 e UTF-32.

A codificação é o processo de transformar um conjunto de caracteres Unicode em uma sequência de bytes. A decodificação é o processo de transformar uma sequência de bytes codificados em um conjunto de caracteres Unicode. A codificação UTF-16 é usada pelo Common Language Runtime para representar os valores Char e String.

Para saber mais acesse <https://docs.microsoft.com/en-us/dotnet/standard/base-types/character-encoding-introduction> e <https://docs.microsoft.com/en-us/dotnet/standard/base-types/character-encoding>.