



Why

We name sequences in $\{0, 1\}$ to easily discuss codes in this set.¹

Definition

A *bit string* (*binary string*) is a finite sequence in the set $\{0, 1\}$.

If a bit string has length one, we refer to it as a *bit*. Using this terminology, it is natural to call the sequence terms *bits*. Other terminology for bit strings includes *binary string*, *bit sequence* and *digital datum*.

If the bit string has length eight, we refer to it as a *byte*. Using this terminology, a *kilobyte* is a length $8 \cdot 2^{10}$ bit string. In other words, a kilobyte is $2^{10} = 1024$ bytes, or roughly one thousand bytes. Likewise a *megabyte* is a length $8 \cdot 2^{20}$ bit string. A megabyte is $2^{20} = 1048576$ bytes, or roughly one million bytes. Similarly a *gigabyte* is 2^{30} bytes and a *terabyte* is 2^{40} bytes.²

Notation

We often denote the set $\{0, 1\}$ by **B**. Using this notation, we denote the length n bit strings by **B**^{*n*}.

¹Why we want to name these requires more explanation, to be included in future editions.

²Warning: some authors use these monikers with powers of ten. For example, a kilobyte is exactly one thousand bytes. etc.

We occasionally use **false** to denote the length-1 bit string $(0,)$ and **true** to denote the length-1 bit string $(1,)$. In this context, **bool** is another name for the set $\{\mathbf{true}, \mathbf{false}\}$ ³.

³Future editions might break this out into its own sheet.

