



Why

1

Definition

Let X be a set and let A be a finite set. We denote the set of all finite sequences (strings) in A by $\mathcal{S}(A)$. We read $\mathcal{S}(A)$ aloud as “the strings in A .” The length zero string is \emptyset .

A *code* for X in A is a function from X to $\mathcal{S}(A)$. In this context, we refer to the finite set A as an *alphabet* and we call $c(x)$ the *codeword* of x . The *length* of $x \in X$, with respect to a code $c : X \rightarrow \mathcal{S}(A)$, is the length of the sequence $c(x)$ (its codeword). We call a code *nonsingular* if it is injective.

Examples

Define $c : \{\alpha, \beta\} \rightarrow \{0, 1\}$ by $c(\alpha) = (0,)$ and $c(\beta) = (1,)$.²

Code extensions

Let $s, t \in \mathcal{S}(A)$ of length m and n respectively. The *concatenation* of s with t is the length $m + n$ string $u \in \mathcal{S}(A)$ defined by $u_1 = s_1, \dots, u_m = s_m$ and $u_{m+1} = t_1, \dots, u_{m+n} = t_n$. We denote the concatenation of s and t by st . Note, however, that $st \neq ts$, although $s(tr) = (st)r$.

¹Future editions will include, with perhaps discussion of encoding and representing text.

²Future editions will include additional examples.

Given a code $c : X \rightarrow \mathcal{S}(A)$, we can produce a code for $\mathcal{S}(X)$ in a natural way. The *extension* of c is the function $C : \mathcal{S}(X) \rightarrow \mathcal{S}(A)$ defined, for $\xi = (\xi_1, \dots, \xi_n) \in \mathcal{S}(X)$, by

$$C(\xi) = c(\xi_1) \cdots c(\xi_n).$$

We call an code *uniquely decodable* if its extension is injective. In other words, given the code $C(\xi)$ for a sequence $\xi \in \mathcal{S}(X)$, we can recover ξ . We call $C(\xi)$ the *encoding* of ξ . We call ξ the *decoding* of $C(\xi)$.

