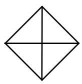


Why

[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."

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*. The *length* of $x \in X$, with respect to a code $c : X \to \mathcal{S}(A)$, is the length of the sequence $c(x)$. We call a code *nonsingular* if it is injective.

Examples

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

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 $s(tr) = (st)r$.

Given a code $c : X \to \mathcal{S}(A)$, we can produce a code for $\mathcal{S}(X)$ in a natural way. The *extension* of c is the function

[1]Future editions will include, with perhaps discussion of encoding a representing text.

$C : \mathcal{S}(X) \to \mathcal{S}(A)$ 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 $C(\xi)$.

Prefix-free codes

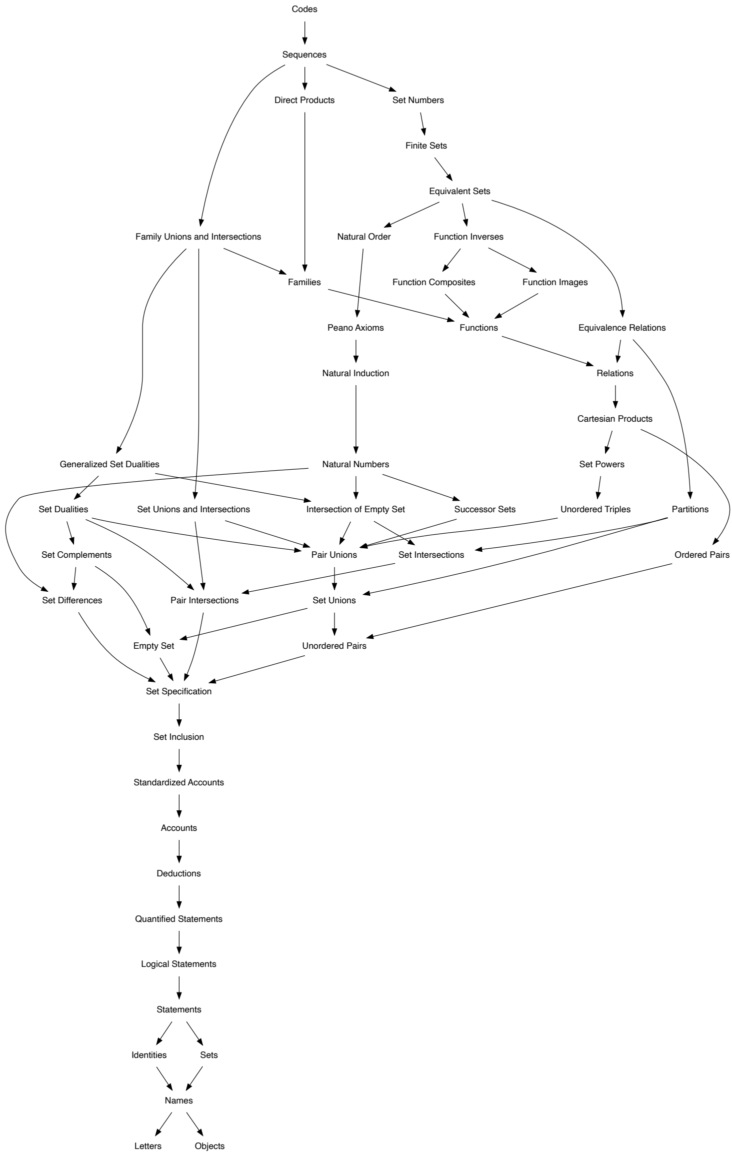
We call a string $s \in \mathcal{S}(A)$ of length m a *prefix* of a string $t \in \mathcal{S}(A)$ of length n if $m \leq n$ and $s_i = t_i$ for all $i \in \{1, 2, \dots, m\}$.

We call a code $c : X \to \mathcal{S}(A)$ *prefix-free* if, for all $x \in X$, $c(x)$ is *not* a prefix of $c(x')$ for all $x' \neq x$, $x' \in X$. Otherwise, we call the code *prefixed*. All prefix-free codes are uniquely decodable, but the converse is false.

Proposition 1. *There exists a set X, alphabet A, and prefixed code $C : X \to \mathcal{A}$ such that C is uniquely decodable.*

Proof. Let α and β be objects. Try $X = \{\alpha, \beta\}$, $A = \{0, 1\}$ and $c : X \to \mathcal{S}(A)$ defined by $c(\alpha) = (0,)$, $c(\beta) = (0, 1)$. We proceed by induction on the length of encodings. Consider a length one encoding. It must be $(0,)$, which decodes as $(A,)$. Consider a length two encoding. It is either $(0, 0)$, which decodes as (A, A), or it is $(0, 1)$ which decodes as $(B,)$. Now assume the cases $k - 1$ and $k - 2$. Now consider a length k code $a \in \mathcal{S}(A)$. It consists of $(a_{1:k-1}, a_k)$. If $a_k = 0$, then the the code must be (y, α) where y is the decoding of $a_{1:k-1}$. By the induction hypothesis, $a_{1:k-1}$ is of length $k-1$ and so uniquely decodable. Otherwise, $(a_{k-1}, a_k) = (0, 1)$ and so the code must be (y', β) where y' is the decoding of $a_{a:k-2}$. By the induction hypothesis, $a_{1:k-2}$ is of length $k-2$ and so uniquely decodable. $\square$

In other words, the prefix-free codes are a *strict* subset of the uniquely decodable codes.


Codes
Sequences
Direct Products
Set Numbers
Finite Sets
Equivalent Sets
Family Unions and Intersections
Natural Order
Function Inverses
Families
Function Composites
Function Images
Peano Axioms
Functions
Equivalence Relations
Natural Induction
Relations
Cartesian Products
Generalized Set Dualities
Natural Numbers
Set Powers
Set Dualities
Set Unions and Intersections
Intersection of Empty Set
Successor Sets
Unordered Triples
Partitions
Set Complements
Pair Unions
Set Intersections
Ordered Pairs
Set Differences
Pair Intersections
Set Unions
Empty Set
Unordered Pairs
Set Specification
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Identities
Sets
Names
Letters
Objects