Why

Let X and A be finite sets. Here are two practical considerations for constructing a code $C : X \to \mathcal{S}(A)$.

First, the code should be nonsingular (injective). In other words, no two objects in the base set have the same codewords.

Second, no additional information should be required to indicate where codewords start and end. This second restriction indicates that no codeword should appear as the first part of another codeword of greater length. This second implication motivates this sheet.

Definition

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 $i \in \{1, 2, \ldots, m\}$.

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

2

$(A,)$. 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. $\qquad\square$

In other words, the prefix-free codes are a *strict* subset of the uniquely decodable codes. However, our second consideration mentioned above indicates that the "practical" codes are prefix-free.

Prefix-Free Codes

Codes

Dot-Dash Code

Flash Code

Alphabets

Lists

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

Set 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

4