



## Why

We return to our discussion of symbols and scripts, to make precise these concepts in the language of sets and lists.

## Definition

An *alphabet* is a finite set. For example, let  $A$  be the set

$$\{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\},$$

where  $a$  denotes the latin lower case letter “a”,  $b$  denotes the latin lower case letter “b”, and so on. In other words,  $A$  is the set of lowercase latin letters. It is an alphabet. By analogy with this familiar case, we frequently refer to the elements of an alphabet as *letters* or *symbols*.

A *word* is a list of letters in an alphabet, and a *phrase* is a list of words. For example,  $(c, a, t, s)$  is a word in  $\mathcal{A}$  (mean to correspond to the word “cats”) and

$$((c, a, t, s), (a, n, d), (d, o, g, s))$$

is a phrase in  $\mathcal{A}$  (meant to correspond to the phrase “cats and dogs”).

## Strings

Let  $A$  be an alphabet. In this case (in which  $A$  is a finite set), we refer to the lists of  $A$  as *strings*. The string whose length is zero is the empty set.

## Notation

We denote the set of all lists (strings) in  $A$  by  $\text{str}(A)$ . We read  $\text{str}(A)$  aloud as “the strings in  $A$ .”

