1 Introduction

Definition 1.1
Let $\Sigma = \{0, \ldots, \sigma - 1\}$ be a finite, ordered set. The elements of Σ are called *characters* or *symbols* and Σ is called an *alphabet* of size σ.

Definition 1.2
A *string* S is a sequence of characters from an alphabet Σ.

- We usually use $n = |S|$ to be the length of the string.
- The i-th character of S is $S[i]$. Indices are 0-based.
- The substring from the i-th to the j-th character is $S[i..j]$.
- A substring with $i = 0$ is called *prefix*. A substring with $j = n - 1$ is called *suffix*.
- The *i-th suffix* is $S[i..n - 1]$.

1.1 Tries

Definition 1.3
Let $S = \{S_0, S_1, \ldots, S_{N-1}\}$ be a set of strings over an alphabet Σ. A *trie* is a tree, where each node represents a different prefix in the set S. The root represents the empty prefix ε. Vertex u representing prefix Y is a child of vertex v representing prefix X, if and only if $Y = Xc$ for some character $c \in \Sigma$. The edge (v, u) is then labeled c.

If S is the set of all suffixes of a string T, the trie is called *suffix trie*.

Example 1.4
Figure 1.1 shows the suffix trie for the string b̈anana$. The dollar sign \$ is a sentinal that does not appear elsewhere in the text. This guarantees, that no suffix is a prefix of another suffix and the suffix trie therefore has $n + 1$ leaves.

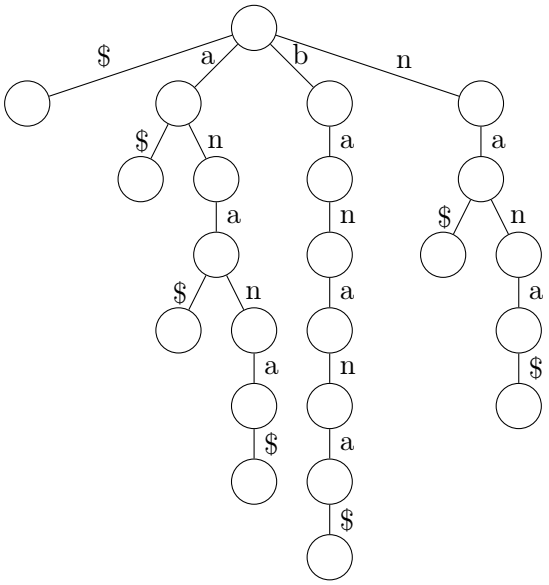

Figure 1.1: Hallo