Data Structures

Worksheet 8

BITL2

25 November 2021

---

1. **Warm up:** Answer the following questions.
   (a) Why is hashing important?
   (b) What is the difference between a map and a hashing function?
   (c) In what cases is a rolling hash function the same as a regular hash function?

2. Draw what happens when the keys $5, 28, 19, 15, 20, 33, 12, 17, 10$, are insterted into a hash table with hash function $h(k) = k \pmod 9$, with collisions resolved by chaining.

3. This question is about **string matching** algorithms.
   (a) Recall the naive string matching algorithm as you saw it in Discrete Structures. Consider the two strings

$$s = \texttt{ambracadambrazampbra}, \qquad t = \texttt{amp}.$$

   How many characters will be compared when t is searched for in s?
   (b) How many of those comparisons for `a` are pointless, because you already know the character is not `a`?
   (c) To fix the problem in part (b), for every string `s` we define the **prefix** function $\pi_{\mathtt{s}} \colon \mathbf{Z}_{\geqslant 0} \to \mathbf{Z}_{\geqslant 0}$, given by

$$\begin{aligned} \pi_{\mathtt{s}}(k) &= \max_{\ell<k} \{\mathtt{s}[:\ell] = \mathtt{s}[k-\ell:k]\} \\ &= \max_{\ell} \{\mathtt{s}[:k][:\ell] = \mathtt{s}[:k][-\ell:]\}. \end{aligned}$$

   Find the values of $\pi_{\mathtt{s}}(k)$ for each $k = 0, \ldots, \mathrm{len}(\mathtt{s})$ for the strings

   i. `grebulon` ii. `aaaaaaaaba` iii. `abaaaaaaaa` iv. `catercatcat`

   (d) Suppose you are given a sequence of nonnegative integers $a_1, \ldots, 1_\ell$. Describe what conditions the sequence must meet to correspond to the values $\pi_{\mathtt{s}}(1), \ldots, \pi_{\mathtt{s}}(\ell)$ of a string `s` of length ℓ. How would you construct `s`?