



Algorithms and Datastructures

Summer Term 2022

Sample Solution Exercise Sheet 12

Due: Wednesday, July 27th, 4pm

Exercise 1: Rabin-Karp Algorithm (10 Points)

- (a) Implement the Rabin-Karp algorithm. You may use the template `StringMatching.py`. The algorithm should return a Python-list containing all starting points of the pattern. That is, for each time the pattern is recognized, the list should contain the position of the first letter of this appearance.
- (b) Run your algorithm on the text and pattern given in `input.txt`. Write the output into `erfahrungen.txt`.
Remark: When choosing the parameters b and M , consider that the procedure `read_input` used on `input.txt` creates an array with values from `ord(' ') = 32` (whitespace) to `ord('z') = 122`.

Sample Solution

- (a) C.f. `StringMatching.py`.
- (b) The desired output is:
`[212, 2194, 2604, 5208, 7193, 7443, 7939, 10245, 11594, 13544, 14276, 22354, 25024, 28735, 39999, 40835, 46199]`.

Exercise 2: Knuth-Morris-Pratt Algorithmus (10 Points)

Consider the pattern $P = BBABAB$ and the text $T = ABBABBABABBABABBA$.

- (a) Compute the array S of the Knuth-Morris-Pratt algorithm. (5 Points)
- (b) Use the Knuth-Morris-Pratt algorithm to find all appearances of P in T . Document the steps analogously to the lecture. (5 Points)

Sample Solution

- (a) $S = [-1, 0, 1, 0, 1, 0, 1]$
- (b)

A	B	B	A	B	B	A	B	A	B	B	A	B	A	B	B	A	
<u>B</u>	B	A	B	A	B												
	B	B	A	B	<u>A</u>	B											
				B	B	A	B	A	B								✓
									B	B	A	B	A	B			✓
														B	B	A	✓