

Digital Libraries

WS 14/15

Jens Metzner & Manuel Wildner

Solution of assignment 1

(Submission: 7. Nov)

Task 1.1: *What are the key tasks of libraries? Describe each task briefly.* [] out of 2 Points

Libraries act as mediators between information and users. They play an important role in information retrieval and knowledge transfer.

- selection - Definition of collections
- acquisition - Physical objects
- description - Catalogs
- access - Shelves, Lending schemes
- preservation - Controlled environment, Media care

Task 1.2: *Give, for each of the tasks, an example for useful application of digital technology.* [] out of 3 Points

Task 1.3: *It is widely believed that information supply fosters economic growth. Give two reasons why this can be the case.* [] out of 2 Points

- Due to the transfer of technology knowledge.
- To see the process of scientific discoveries.

Task 1.4: *Briefly describe the idea how Bush's Memex organizes information and provides access to it.* [] out of 3 Points

Bush's Memex should be a analog device/desk in which individuals would compress and store all of their books, records, and communications in a microfilm storage. Two translucent screens were on the top of the desk on which material could be projected for convenient reading. Buttons were used to switch pages, save documents, turn pages or create 'links' between pages. Such associations trails were shared for example with other scientists.

Task 2.1: *Describe the three main ideas behind the Boyer-Moore algorithm for string searching* [] out of 2 Points

- The pattern is compared against a possible match from right to left.
- 'Bad character shift rule' is used to avoid unnecessary comparisons against a target character.
- 'Good suffix shift rule' is used to align only matching pattern characters against target characters already successfully matched.

Task 2.2: *Outline the steps of exact string matching using Boyer-Moore algorithm in the following case (state which rule is the step following).* [] out of 4 Points

Text: DAVE.GAVE.NANCY.A.BANANA

Pattern: BANANA

BANANA

BANANA

BANANA

using bad character rule

using bad character rule

using bad character rule

finished

Task 2.3: *What is a Trie structure? Describe briefly how it is built and used for string searching.* [] **out of 4 Points**

A trie, also called prefix tree, is a multi-way tree structure. The Root is empty. Every inner node consists of a prefix depending on its level. For each level one letter is added. The leafs contain a whole word each.

When a new element is inserted, every of its prefixes is compared to the corresponding nodes of the current level. The length of the prefix is increased by one for each level. When the prefix can be found, the next prefix is searched in the level below. When a prefix cannot be found, a new node is added in this level containing the current prefix.

If a word is searched in the tree, simply the path has to be followed, comparing each prefix like for insert.