T-301-REIR, REIKNIRIT
HAUST 2018
D5 – STRINGS: SORTING, TRIES, COMPRESSION

Problem 1. *Suppose you need to sort 10 million integers, each in the range 0 to 2^{40}. How would you do it? Which method, among the ones we have seen, gives the smallest tilde time complexity.*

Problem 2. *Describe an input instance for which MSD is much faster than LSD.*

Problem 3. *Construct a set S of strings in lowercase ($R = 26$) for which the R-way trie uses space (in words) less than the total number of characters in strings in S. (The number of characters in the set $\{$`abc, bcd, aabd`$\}$ is 10.)*

Problem 4. *Give the LZW encoding the string a^N consisting of N repeats of the character `a`. What is the compression ratio as a function of N?*

Problem 5. *(Problem 5.5.18) Let F_k be the k-th Fibonacci number. Consider N symbols, where the k-th symbol has frequency F_k. Note that $F_1 + F_2 + \ldots F_N = F_{N+2} - 1$. Describe the Huffman code. (Hint: The longest codeword has length $N - 1$).*

Problem 6. *Problem IX (Ternary Search Tries) in the final exam 2015.*

Class Exercises

These questions will be addressed during exercise class. They are not to be turned in.

Problem 7. *Suppose you have a trie with $R = 26$ (lowercase letters) and insert into it the words: `tic, tac, toe, time`. How many null links will be stored in the nodes of the trie?*

Problem 8. *(Final Exam 2017) Decode the following LZW-compressed message:*

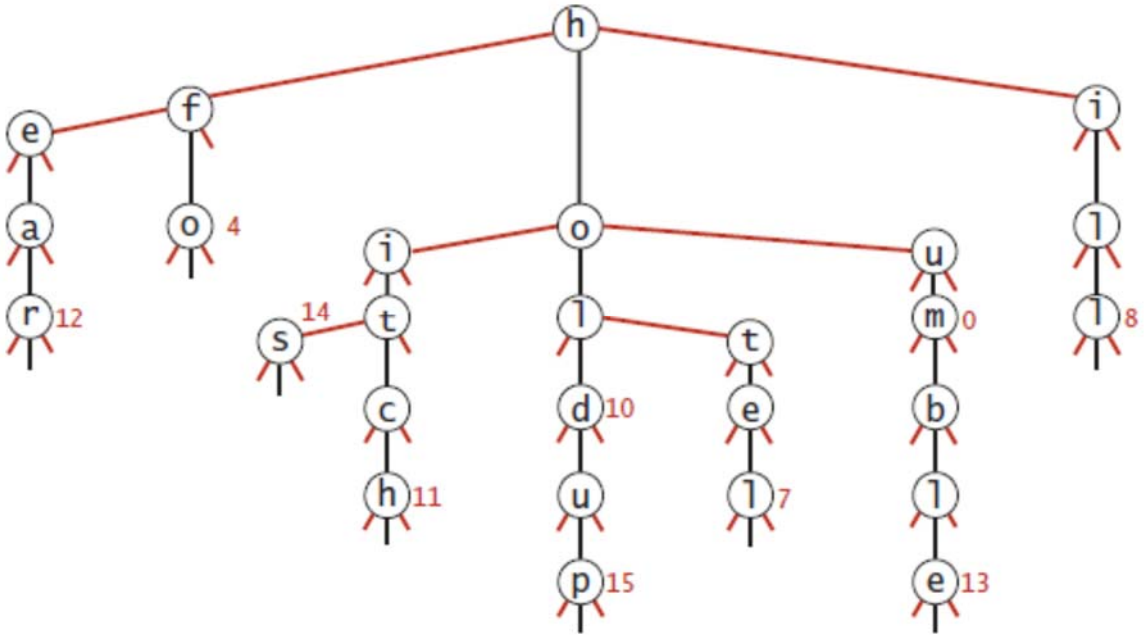
```
41 81 42 82 83 80
```

Problem 9. *Encode the string "ABABCABCD" using LZW.*

Problem 10. *Suppose we are given the following letter frequencies: A:60, D:6, E:75, F:3, G:9, H:4, S:17. Give the optimal Huffman coding tree. Also, how is "SEEAHEAD" compressed?*

XI. (8%) Ternary Search Tries

Að neðan er útkoma sem fæst úr því að setja safn af strengjum (ásamt tengdum heiltölugildum) inn í þríleitartræ (ternary search trie). / *Below is the result of inserting a set of strings (and associated integer values) into a ternary search trie.*



(a) Teldu upp í stafrófsröð strengina sem voru settir inn. / *List in alphabetical order the set of strings that were inserted.*

Svar: _____

(b) Bættu við strengnum hobo, með gildið 18, og strengnum hung, með gildið 21, inn í træið, og sýndu breytingarnar á træinu á myndinni að ofan.
/ *Add the string* hobo *(with associated value 18) and the string* hung *(with associated value 21) and the string, and illustrate the changes to the trie in the figure above.*