University of Wrocław: Data Science 08.07.2022

Math and coding problems

NOTE. You should send us a short report (.pdf file) and programs you used:

- For Tasks 1, 2 and 5 provide numerical answers in a report and attach your programs you used (you may use any programming language).
- For Tasks 3 and 4 provide answers in a report file.

Task 1. Consider numbers written in the decimal system. For set of digits D we say that the number is D-contained if all its digits belong to D. For instance 112, 12333 are $\{1,2,3\}$-contained, while 14 is not.
Write a program which checks how many k-digit $\{8,9\}$-contained prime numbers are there. Present the answers for $k = 5$, and $k = 10$. The answer should be found in a few seconds (for both cases).

Task 2. We say that the sequence of words $S = (w_1, \ldots, w_n)$ is a *partition* of w if $w = w_1 w_2 \ldots w_n$ (i.e. w is a concatenation of all words in S). The *value* of the sequence of words is a sum of squared lengths of words in S, i.e. $\text{value}(S) = \sum_{i=1}^{n} |w_i|^2$. Write a function which for a given text w and a set of words V finds a partition S of w, equal to $(w_1, \ldots, w_n)$ such that:

- each w_i belongs to V
- value(S) is maximal

Your program should write both the partition, and its value. Use the following set of words:

> *h ca go lit they was ant there in now pencil sand licked silly do us sha what love when m real at itwhen say er nearly not of rated out man writing fell wanted ran soft hb any here we hey want ear ease knife no himself whatnot anything ruling bear it him clicked marked blue an earth this ruler oo pooh their to wo ee saying line v his me business if how wrong lo he pin o these or special as words own thing on help walk nothing india w had rubbing hat inch inches p thin near please were is pencils her for be stuff helping sharpening many pockets over who b red all pleased a and case wash know things hi mark i spelt saw word eeyore shut poohs you set ing brave sw little r ask ho bb ow the early with click green ate so down which pocket rubber lovely ha lick she lines but oh tin pen*

Test your program on the following texts:

Text A

```
buteeyorewassayingtohimselfthiswritingbusinesspencilsandwhatnotove
rratedifyouaskmesillystuffnothinginit
```

Text B

```
whenpoohsawwhatitwashenearlyfelldownhewassopleaseditwasaspecialpen
cilcasetherewerepencilsinitmarkedbforbearandpencilsmarkedhbforhelp
ingbearandpencilsmarkedbbforbravebeartherewasaknifeforsharpeningth
epencilsandindiarubberforrubbingoutanythingwhichyouhadspeltwrongan
darulerforrulinglinesforthewordstowalkonandinchesmarkedontheruleri
ncaseyouwantedtoknowhowmanyinchesanythingwasandbluepencilsandredpe
ncilsandgreenpencilsforsayingspecialthingsinblueandredandgreenanda
lltheselovelythingswereinlittlepocketsoftheirowninaspecialcasewhic
hshutwithaclickwhenyouclickeditandtheywereallforpooh
```

Important: Before pasting these texts into your program please delete new lines – neither of these texts should contain any whitespace characters (best option: use files sent separately).
Your program should print the result in a few seconds!

Task 3. A random vector (X, Y) has a continuous distribution with a density function

$$f(x,y) = \begin{cases} c \cdot x & \text{for } \ 0 \le x \le 2, \quad \max(0, 1-x) \le y \le 2-x, \\ \\ 0 & \text{otherwise.} \end{cases}$$

where $c > 0$ is a constant. Compute the variance of a Y conditioned on $X = 1.5$, i.e., $\mathrm{Var}(Y|X = 1.5)$.

Task 4. Let $X_1, \ldots, X_{10}$ be the independent identically distributed random variables coming from a population with a uniform distribution $U(0, \theta)$, $\theta \in (0, \infty)$. Find the likelihood ratio test in the problem of verifying $H_0 : \theta = 1$ against $H_1 : \theta \neq 1$. What is the power of the test under the alternative $\theta = 2$? The significance level $\alpha = 0.1$.

Task 5. We have 24 squares with sides 1, 2, 3, ..., 24cm. Write a program that places some of them (maybe all) on a square board of side 70cm in such a way that the following conditions are met:

a) the coordinates of the upper left corner of each square are integers,

b) each square fits entirely on the board,

c) the squares do not overlap (but may meet, with corners or sides),

d) the sides of the squares to be arranged are parallel to the corresponding sides of the board.

Try to minimize area of the board that remains uncovered. Your program should work less than a minute.
The proposed solution should be printed in the following format:

- the first line should contain the area of the board that remains uncovered,
- it should be followed by 70 lines of 70 characters representing the arrangement of squares, each character is a dot (meaning no square) or the letter 'ABCDEFGHIJKLMNOPQRSTUVWX' ('A' means belonging to a square with side 1, 'B' – with side 2, etc.).