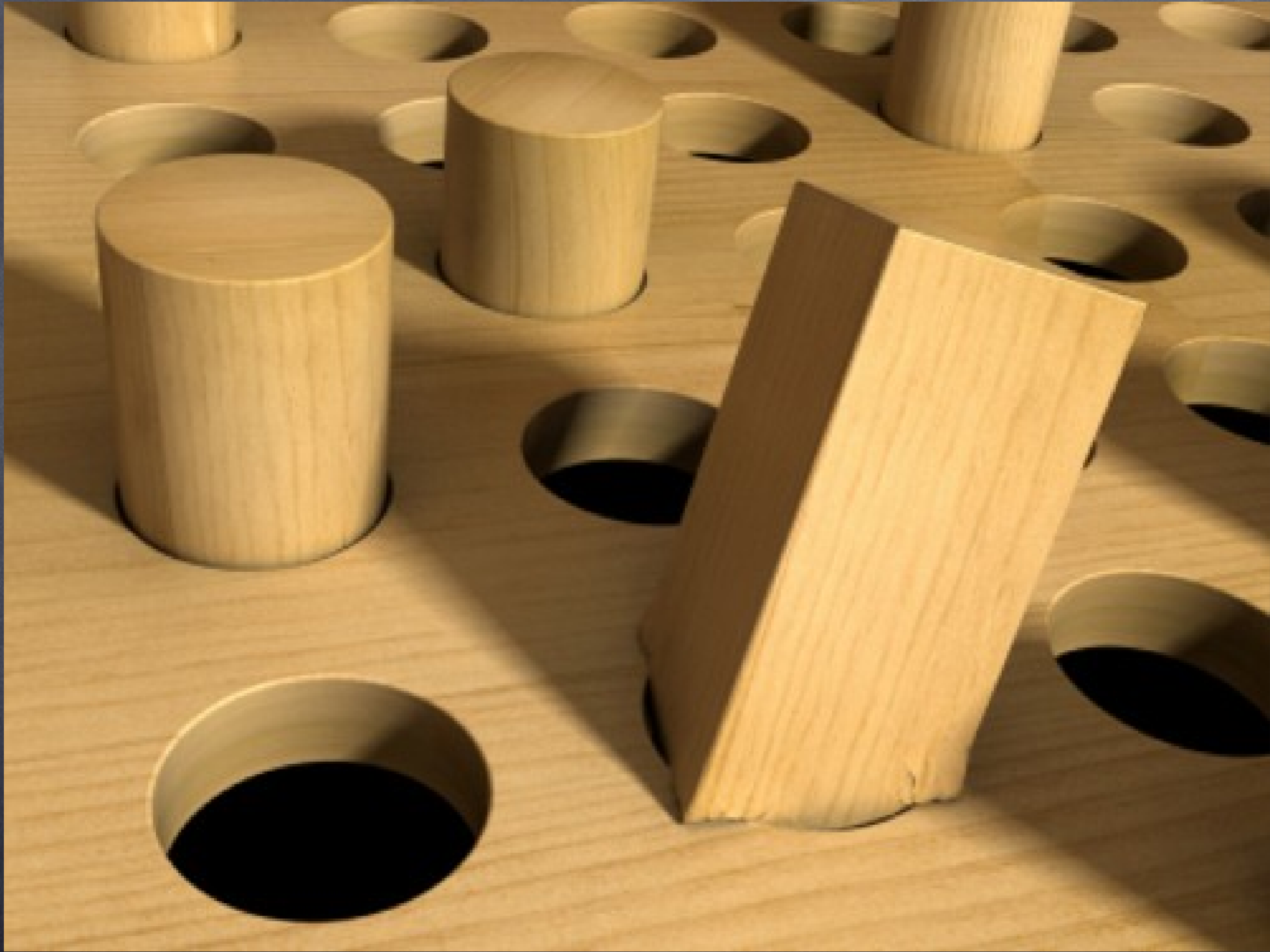


Регулярни изрази

магия в действие

Какво е регулярен израз?

- regular expression | regexp | regex | re
- черна магия
- криптиран запис
- средство за мъчение



Regexpс He ca camo B java

- grep

- sed

- awk

- vim

- emacs

- idea

- perl

- python

- ruby

- javascript

- .net

Търсене и замяна

- `"java".matches("java")` // true
- `"java".replaceFirst("a", "A")` // jAva
- `"java".replaceAll("a", "A")` // jAvA

Специални (мета) СИМВОЛИ

- . | [] {} () ^ \$? * + \
- \ - неутрализира специални символи

Символни класове

- `[]`
- `"j".matches("[a-z]")`
- `"7".matches("[^a-z]")`
- обединение - `[a-d[m-p]]`
- сечение - `[a-z&&[def]]`
- изваждане - `[a-z&&[^m-p]]`

Предефинирани СИМВОЛНИ КЛАСОВЕ

- `.` - Произволен символ
- `\d` - число
- `\D` - $[\text{^}\d]$
- `\s` - интервал, табулация, нов ред
- `\S` - $[\text{^}\s]$
- `\w` - $[\text{a-zA-Z_0-9}]$
- `\W` - $[\text{^}\w]$

Повторения

- $X?$ - нула или едно
- X^* - нула или повече
- X^+ - повече
- $X\{n\}$ - точно n
- $X\{n, \}$ - най-малко n
- $X\{n, m\}$ - между n и m

Групиране

- `()`
- `(one|two|three)` - алтерниране
- `(?<name>X)` - именована група
- `\1` - референция
- `\k<name>` - именована референция
- `(?:X)` - групиране без прихващане

Pattern class

- `boolean Pattern.matches(String, String)`
 - бърза проверка
- `Pattern Pattern.compile(String)`
- `Matcher p.matcher()`
- `String[] p.split()`

Matcher

- `m.find()` - търси съвпадение на шаблон
- `m.start()` - индекс на първия символ
- `m.end()` - индекс на полето след последния символ
- `m.group()` - низ със съвпадението
- `m.replaceFirst()`
- `m.replaceAll()`

КОТВИ

\wedge - начало на ред

$\$$ - край на ред

$\backslash b$ - начало или край на дума

$\backslash B$ - не е граница на дума

$\backslash A$ - начало на низ

$\backslash G$ - край на предишно търсене

$\backslash Z$ - край на низ, преди последния нов ред

$\backslash z$ - край на низ

Упражнение 1

• Напишете програма, която вади количествена статистика за даден текст:

- препинателни знаци
- букви
- думи
- изречения
- параграфи

Следва продължение...