

**Задача за курсов проект:** имплементация на алгоритъм за построяване по зададена дума  $w \in \Sigma^*$  на минимален детерминиран ацикличен автомат  $\mathcal{A}_w$  с език множеството от суфиксите на  $w$ .

За да участвате в надпреварата:

1. до 10.01.2021 г. включително изпратете на pmitankin@fmi.uni-sofia.bg и на petar@lml.bas.bg писмо със subject BASD-66666, където 66666 е Вашият факултетен номер;
2. текстът на писмото задължително да съдържа Вашите имена;
3. към писмото да има прикрепен zip файл 66666.zip, където 66666 е Вашият факултетен номер;
4. файлът 66666.zip да съдържа директория 66666, където 66666 е Вашият факултетен номер;
5. директорията 66666 да съдържа файл 66666.sh
6. за unzip-ването на 66666.zip да се изисква следната парола:  
basd

Вашият курсов проект ще бъде тестван така:

1. файлът 66666.zip ще бъде unzip-нат;
2. cd 66666
3. time sh 66666.sh име-на-входен-файл

Входният файл представлява думата  $w$  и съдържа около 200 000 000 символа. Всеки символ във файла е с размер един байт и представлява цяло число от  $\Sigma = \{i \in \mathbb{Z} \mid 97 \leq i \leq 122\}$ . Входният файл не съдържа символи, които не са от  $\Sigma$ .

**Изисквания за напълно успешен проект:**

1. sh 666666.sh да свърши за не повече от 20 минути върху зададения входен файл;
2. да отпечата в конзолата следните 4 числа:
  - броя на състоянията в автомата  $\mathcal{A}_w$ , т.е.  $|Q_w|$ , където  $Q_w$  е множеството от състоянията на автомата  $\mathcal{A}_w$ ;
  - броя на преходите в автомата  $\mathcal{A}_w$ , т.е.  $|\{(q, x) \in Q_w \times \Sigma \mid \delta_w(q, x) \text{ е дефинирано}\}|$ , където  $\delta_w : Q_w \times \Sigma \rightarrow Q_w$  е частичната функция на преходите в  $\mathcal{A}_w$ ;
  - броя на финалните състояния в автомата  $\mathcal{A}_w$ , т.е.  $|F_w|$ , където  $F_w$  е множеството от финалните състояния на автомата  $\mathcal{A}_w$ ,
  - $|\{v \in \Sigma^* \mid v \text{ е най-дълъг представител} \wedge (\exists u \in \Sigma^*)(v = uu)\}|$ , като  $v$  е най-дълъг представител ако и само ако  $(v \text{ е префикс на } w) \text{ или}$   
(за някои символи  $a, b \in \Sigma$ , такива че  $a \neq b$ ,  
е изпълнено, че  $av$  и  $bv$  са инфикси на  $w$ ).

**Среда.** Вашият проект ще бъде тестван на машина с операционна система Ubuntu 20.04.1, която може да се свали оттук:

<http://www.lml.bas.bg/~petar/ubuntu-20.04.1-desktop-amd64.iso>

или оттук:

<https://releases.ubuntu.com/20.04/ubuntu-20.04.1-desktop-amd64.iso>

(и да се инсталира на виртуална машина напр.) На машината ще има поне 16 GB свободна памет RAM.

На машината ще има инсталирана java, като `java -version` връща

```
java version "1.8.0_261"  
Java(TM) SE Runtime Environment (build 1.8.0_261-b12)  
Java HotSpot(TM) 64-Bit Server VM (build 25.261-b12, mixed mode)
```

За да `zip`-нете директория `66666`, където `66666` е Вашият факултетен номер, с парола `basd`

може да използвате следната команда:

```
zip -er 66666.zip 66666
```

където `66666` е Вашият факултетен номер.

**Внимание.** Ако файлът `66666.zip`, който изпращате на `pmitankin@fmi.uni-sofia.bg` и на `petar@lml.bas.bg`, където `66666` е Вашият факултетен номер, не е `zip`-нат с парола, има голям риск `mail-server`-ите да разпознаят, че `zip` файлът, който изпращате, съдържа изпълним код и поради тази причина изпратеното от Вас писмо да не пристигне нито на `pmitankin@fmi.uni-sofia.bg`, нито на `petar@lml.bas.bg`.

**Други условия.** Вашият проект няма да бъде компилиран преди извикване на `sh.666666.sh`. Очаква се, че ще работи при извикване на `sh 666666.sh`.

Всеки проект трябва да бъде защитен от автора си пред преподавателите. Незащитен проект се оценява с 0 точки. Ако Вашият проект не отговаря на нито едно от изискванията за напълно успешен проект (виж по-горе), но Вие го защитите успешно, ще получите 0.5 точки. Ако Вашият проект отговаря на всички изисквания за напълно успешен проект и Вие го защитите успешно, ще получите 1 точка<sup>1</sup>. Онзи, чийто проект е напълно успешен и дава най-добро `user + sys` време и си защити проекта, получава директно оценка отличен 6 като крайна оценка за курса. Ако най-доброто `user + sys` време на напълно успешен проект е  $X$ , а следващото време на напълно успешен проект е  $Y$  и  $Y < X + X/20$ , то се приема, че  $X=Y$  и в този случай няма най-добро време и никой не получава директно оценка отличен 6.

Проекти, които не могат да бъдат изтествани дали са успешни или не на описаната по-горе среда, напр. такива, написани на `C#`, не могат да участват в надпреварата за директна оценка отличен 6, но може да бъдат защитени и изтествани дали са успешни или не по време на устен изпит през сесията, като за целта студентите

---

<sup>1</sup>при точно едно вярно отпечатано число получавате 0.666 точки, а при точно две или точно три вярно отпечатани числа получавате 0.833 точки.

трябва да имат на разположение подходящи машини за изтестване на проекта по време на изпита.

Проекти изпратени след 10.01.2021 и преди ( $\leq$ ) 21.01.2021 не участват в надпреварата за директна оценка отличен 6, но ще бъдат изтествани дали са успешни преди изпита.

Проекти изпратени след ( $>$ ) 21.01.2021 няма да бъдат тествани дали са успешни преди изпита.

Желаем Ви успех!