

# Regulární výrazy – seznam úloh

---

Seznam úloh slouží k procvičení práce s regulárními výrazy.

## 1 Kontrola řetězců

### 1.1 Úloha

Otestujte, zda se v zadaném slově nachází písmeno „a“.

### 1.2 Úloha

Definujte regulární výraz pro otestování, zda text obsahuje slovo „příloha“, „přílohy“, „příloze“, „přílohu“ nebo „přílohou“.

### 1.3 Úloha

Definujte regulární výraz pro test, zda řetězec obsahuje pouze písmena (malá/velká)

### 1.4 Úloha

Definujte regulární výraz pro test, zda je řetězec věta – věta začíná velkým písmenem a na konci je tečka. Tečku uvnitř věty neuvažujte (slova např., viz., apod., ...).

Úlohu modifikujte tak, že věta může být zakončena i znaky „?“ a „!“.

### 1.5 Úloha

Napište regulární výraz pro ověření elementárního matematického příkladu ve tvaru „5 + 3 = 8“.

### 1.6 Úloha

Napište regulární výraz pro test, zda je řetězec ve formátu:

- a) CamelCase – jednotlivá slova začínají velkým písmenem a ostatní písmena jsou malá
- b) snake\_case – jednotlivá slova jsou oddělena podtržítkem, ostatní písmena jsou malá

### 1.7 Úloha

Definujte regulární výraz pro kontrolu adresy. Adresa je ve tvaru „Zelená 523/15“, název ulice však může být víceslovný („Pod červeným hájem“).

### 1.8 Úloha

Definujte regulární výraz pro kontrolu rodného čísla ve tvaru (521123/0336). Pokuste se vzít v úvahu maximální možné hodnoty měsíce a dne a ženské pohlaví.

### 1.9 Úloha

Napište regulární výraz pro kontrolu PSČ. PSČ je ve tvaru „130 00“.

### 1.10 Úloha

Definujte regulární výraz pro kontrolu emailové adresy. Adresa může být ve tvaru „kacenska18@seznam.cz“, ale i „kotelnik@fel.cvut.cz“

### 1.11 Úloha

Definujte regulární výraz pro kontrolu telefonního čísla v mezinárodním formátu +420608123456.

Úlohu modifikujte tak, aby bylo číslo platné i s mezerami mezi trojicemi – „+420 608 123 456“

### 1.12 Úloha

Definujte regulární výraz pro kontrolu SPZ ve tvaru 3A5 3428, SPZ může zároveň být i ve druhém tvaru 1AA 1598.

### 1.13 Úloha

Pomocí regulárního výrazu rozhodněte, zda zadaný text obsahuje HTML znázornění tučného písma nebo kurzívy. Pro tučné písmo musí být text obalen elementy `<strong>tucne pismo</strong>` a pro kurzivu `<em>kurziva</em>`.

### 1.14 Úloha

Zkontrolujte pomocí regulárního výrazu platnost webové adresy. Formát URL je

`http://www.google.com` (tzn. definice protokolu, znaky „:“ a „/“ a doména). Nezapomeňte, že adresa může být i `http://www.fel.cvut.cz`

Úlohu modifikujte na pokročilou kontrolu URL, tzn., že může obsahovat plnou cestu a Query String – např. `http://www.test.cz/slozka/podslozka/abc.html?parametr1=hodnota1&parametr2=hodnota2`

### 1.15 Úloha

Definujte regulární výraz pro kontrolu adresy IPv4. Formát adresy je 192.168.0.1.

Úlohu modifikujte tak, aby IP adresa byla ve formátu CIDR (192.168.0.1/24) – obsahuje lomítko a za ním počet jedniček v masce.

### 1.16 Úloha

Definujte regulární výraz pro kontrolu MAC adresy. MAC adresa má 6B zapsaných hexadecimálně.

Jednotlivé byty jsou odděleny znakem „:“, „-“ nebo nemusí být vůbec odděleny. Každé hexadecimální číslo má právě dva znaky. Příkladem je FF:8A:32:07:15:01, FF-8A-32-07-15-01 nebo FF8A32071501

### 1.17 Úloha

Definujte regulární výraz pro kontrolu adresy IPv6. Formáty adres mohou být následující:

- 2001:0db8:0000:0000:0000:0000:1428:57ab
- 2001:0db8:0000:0000:0000::1428:57ab
- 2001:0db8:0:0:0:0:1428:57ab
- 2001:0db8:0:0::1428:57ab
- 2001:0db8::1428:57ab
- 2001:db8::1428:57ab

Znamená to, že adresa obsahuje 8 částí, obsahující hexadecimální čísla, oddělených dvojtečkou.

V případě, že by daná část obsahovala pouze nuly, je možné tuto část nahradit jednou nulou.

Libovolný počet částí po sobě obsahující pouze nuly je možné nahradit čtyřtečkou.

## 2 Nahrazování řetězců

### 2.1 Úloha

Nahradte v textu více mezer za sebou jedinou (text mnohdy obsahuje dvě a více mezer jdoucích po sobě).

### 2.2 Úloha

Očistěte text od párových HTML značek. Párová HTML značka je vždy ve tvaru `<znacka>text</znacka>`.