# Упражнения: Регулярни изрази

Тествайте решението в Judge: <https://judge.softuni.bg/Contests/3177/Regular-Expressions>.

## Търсене на пълно име

Напишете програма на C#, която търси всички пълни имена от списък с имена. Накрая ги отпечатайте.

### Писане на регулярни изрази

Първо трябва да напишем **регулярен израз** с валидни имена по следните изизквания:

* Пълното име е валидно, когато
  + Съдържа само две думи
  + Всяка дума започва с главна буква
  + След първата буква всички останали букви в думата са малки
  + Всяка дума трябва да има минимална дължина от 2 символа.
  + Между двете букви трябва да има интервал

За да ви помогнем, ще ви покажим няколко стъпки:

1. Използвайте онлайн сайт за тестване на регекс като [https://regex101.com](https://regex101.com/)
2. Вижте как да използвате набори от символи (обозначени с квадратни скоби - "[]")
3. Уточнете че търсете две думи с интервал между тях (приема се само символа **'** **',** без табове и нови редове)
4. За всяка дума уточнете, че трябва да започва с главна буква от английската азбука.
5. Останалите букви след първата трябва да бъдат малки. Ще има една или повече малки букви – използвайте правилния **quantifier**.
6. За да предотвратите засичането на букви на нови редове, сложете "\b" в началото и края на вашия регекс. .

С цел да проверите вашия регекс може да използвате таблицата по-долу.

|  |  |
| --- | --- |
| **Верни изрази** | **Грешни изрази** |
| Ivan Ivanov | ivan ivanov, Ivan ivanov, ivan Ivanov, IVan Ivanov, Ivan IvAnov, Ivan Ivanov |

Резултатите от търсенето трябва да изглеждат така:

Graphical user interface, application

Description automatically generated

След като сте създали регулярния израз трябва да го напишете в C#

### Имплементация на решението в C#

Създайте нов проект на C# и копирайте регулярния израз в **низ**:

A picture containing table

Description automatically generated

Забележка: Добра идея е да използвате символа '@' за да складирате регулярния израз, защото някои символи като '\' могат да приченят грешка.

Време е да **прочетем входа** и да **извлечем съвпаденията**. Затова ще използваме класа MatchCollection

Text

Description automatically generated with medium confidence

След това извличаме резултатите, минаваме през **MatchCollection** и принтираме намереното

Graphical user interface, text, application

Description automatically generated

### Примери

|  |
| --- |
| **Вход** |
| Ivan Ivanov, Ivan ivanov, ivan Ivanov, IVan Ivanov, Test Testov, Ivan Ivanov |
| **Изход** |
| Ivan Ivanov  Test Testov |

## Търсене на телефонни номера

Напишете регулярен израз, който да търси валидни номера от София. След като ги намерите, трябва да ги отпечатате на конзолата, разделени със запетая и интервал “, ”.

### Писане на регулярни изрази

Номера е валиден, когато са изпълнени следните условия:

* Започва с "**+359**"
* След това трябва да има винаги 2
* Следващите числа трябва да бъдат 7 наброй (разделени на две групи от 3 и 4 числа):
* Различни части трябва да бъдат разделени с интервал или тире ('**-**').
* Използваме **quantifier-и** за да намерим специфичните числа.
* Използвайте група, за да се уверите, че разделителят е един от следните символи: интервал или тире. Не се приема комбинация от двата символа (+359 2-111 111 е невалиден, защото съдържа тире и интервал). Използвайте **обратна референция.**Добавете граница на думата в края на съвпадението, за да избегнете частични съвпадения (последният пример от дясната страна).
* Уверете се, че знака '+' е в **начало на низа**.

Можете да използвате таблицата по-долу за да си проверите регекса:

|  |  |
| --- | --- |
| **Верни изрази** | **Грешни изрази** |
| +359 2 222 2222  +359-2-222-2222 | 359-2-222-2222, +359/2/222/2222, +359-2 222 2222  +359 2-222-2222, +359-2-222-222, +359-2-222-22222 |

### Имплементация на решението в C#

Сега е време да напишем решението, затова да запредваме ръкави!

Първо трябва да създадем променлива както в предишната задача:

A picture containing text

Description automatically generated

Отново използваме класа MatchCollection за търсене на съвпадения

Text

Description automatically generated

Нека да се опитаме да отпечатаме съвпаденията само чрез един ред код. Понеже MatchCollection връща колекция, може да използваме метода LINQ.

За да вземем всички съвпадения в масив от низ трябва да изпълним няколко манипилации в MatchCollection:

1. Превръщаме всеки един елемент от MatchCollection в тип Match чрез Cast<Match>().
2. Тъй като всеки елемент е Match, можем да вземем стойностите на всяко съвпадение чрез Select(). Може да използваме метода Trim() за стойностите, за да премахнем всички начални или крайни интервали
3. След като взехме стойностите може да използваме ToArray(), за да превърнем колекцията в масив

Ето така изглежда нашата LINQ заявка:

Text

Description automatically generated with medium confidence

След това отпечатваме валидните номера чрез string.Join():

Text

Description automatically generated with medium confidence

### Примери

|  |
| --- |
| **Вход** |
| +359 2 222 2222,359-2-222-2222, +359/2/222/2222, +359-2 222 2222 +359 2-222-2222, +359-2-222-222, +359-2-222-22222 +359-2-222-2222 |
| **Изход** |
| +359 2 222 2222, +359-2-222-2222 |

## Търсене на дати

Напишете програма, която търси дати във формата “dd{разделител}MMM{разделител}yyyy”. Използвайте **групи** в регулярния израз.

### Писане на регулярни изрази

Датата е валидна, когато са изпълнени следните условия:

* Винаги започва с две числа последвани с раздетел.
* Има една главна и след това две малки букви. (Jan, Mar).
* След това има разделител последван от точно четири числа (за годината).
* Разделителя може да бъде три типа: точка (“.”), тире (“-“) и наклонена черта (“/”).
* Разделителя трябва да раздели цялата дата например (13.03.2016 е валидно, 13.03/2016 е невалидно). Използвайте **обратна референция,** за да проверите това.

Може да следвате таблицата по-долу за регекса:

|  |  |
| --- | --- |
| **Верни изрази** | **Грешни изрази** |
| 13/Jul/1928, 10-Nov-1934, 25.Dec.1937 | 01/Jan-1951, 23/sept/1973, 1/Feb/2016 |

Използвайте **групи** за **деня**, **месеца**, и **годината.**

Тъй като задачата изисква по-сложен регекс, включва наименувани групи. Нека сега да видим как се създава:

* Първо не искаме да има други символи преди нашата дата, затова ще използваме “\b”:  
  Graphical user interface, application

  Description automatically generated
* След това ще търсим деня. Уточняваме, че търсим точно две цифри и тъй като по-късно ще искаме да извлечем този ден, ще го сложим в група:  
  Table

  Description automatically generated with low confidence
* Ще добвим име на групата, за да я търсим по-лесно:  
  Table

  Description automatically generated with low confidence
* След това идва разделителя - тире, точка или наклонена черта. Може да използваме клас от символи:  
  A picture containing table

  Description automatically generated
* Понеже искаме да използваме същия разделите и по-късно, можем да го сложем в група:  
  A picture containing graphical user interface

  Description automatically generated
* След това идва месеца, който съдържа главна латинска буква и точно две малки латински букви:  
  A picture containing application

  Description automatically generated
* След това ще търсим предишния разделител. Може да използваме обратна референция за това:  
  A picture containing application

  Description automatically generated
* Сега ще търсим годината, която се състои от точно четири цифри:  
  A picture containing application

  Description automatically generated
* Тъй като не искаме да има други символи след датата можем отново да използваме “\b”:  
  A picture containing table

  Description automatically generated

Време е да намерим **валидните дати** и да ги отпечатаме на входа на **нов ред** във формата: “Day: {ден}, Month: {седмица}, Year: {година}”,

### Имплементация на решението в C#

Първо ще сложим регекса в променлива и ще получим MatchCollection:

Graphical user interface

Description automatically generated with medium confidence

Тъй като регекса работи различно в различните езици, преди да продължим, ще променим обратната референция от \2 **на** \1 Правим така, защото обратните референции в C# не броят наименуваните групи за обратни референции.

След това минаве през всеки един Match и извличаме деня, месеца и годината от групите. Може да използваме специален синтаксис в C#, който да вземе стойностите като ключ, както взимаме стойностите от речник:

Text

Description automatically generated

### Примери

|  |
| --- |
| **Вход** |
| 13/Jul/1928, 10-Nov-1934, , 01/Jan-1951,f 25.Dec.1937 23/09/1973, 1/Feb/2016 |
| **Изход** |
| Day: 13, Month: Jul, Year: 1928  Day: 10, Month: Nov, Year: 1934  Day: 25, Month: Dec, Year: 1937 |

## Мебели

Напишете програма, която пресмята цената на всички мебели. Ще получавате предмети докато не получите **"Purchase"**. Всеки ред ще бъде в следния формат:

**">>{мебел}<<{цена}!{качество}"**

Цената може да бъде цяло число или число с плаваща запетая. Отпечатайте мебелите и общата цена. На всеки ред трябва да отпечатате предеметите в следния формат:

**"Bought furniture:**

**{мебел1}**

**{мебел2}**

**…"**

Накрая трябва да отпечатате следното съобщение: "Total money spent: {общата цена}" **-** форматирайте сумата до второто число след десетичната запетая.

Ако получите невалиден вход, пропуснете командата.

### Примери

|  |  |  |
| --- | --- | --- |
| **Вход** | **Изход** | **Коментари** |
| >>Sofa<<312.23!3  >>TV<<300!5  >Invalid<<!5  Purchase | Bought furniture:  Sofa  TV  Total money spent: 2436.69 | Само **Sofa и TV** са валидни. Умножаме цените им по количеството и отпечатваме разултата |

## Приходите на бара в СофтУни

Нека да си вземем почивка и да отидем в бара на СофтУни. Време е хората от бара да се прибират по домовете си. Вие сте човекът, който трябва да пресметни оборота за деня. Докато не получите "end of shift", ще получавате редове от команди, но преди да обработите редовете, ще трябва да направите няколко проверки.

Всеки правилен вход трябва да има **име на** **клиент**, **продукт**, **количество** и **цена**:

* Ако името на клиента е **заобградено с** **"%"** и започва с главна буква, следвана от малки, то той е валиден
* Ако един продукт **съдържа букви** и е **заобграден с** **"<"** и **">"**, то той е валиден
* Ако количеството е **число**, **заобградено с "|"**, то то е валидно
* Ако цената е **реално число**, **последвано с** **"$"**, то тя е валидна

Всеки валиден ред трябва да бъде в реда: **име на клиент, продукт, количество и цена**.

Между всяка част може да има и други символи, с изключение на (**'|', '$', '%' и '.'**)

За всеки валиден ред трябва да от печатате на конзолата:

"{име на клиент}: {продукт} - {оборот}"

Когато получите "end of shift", отпечатайте оборота за деня, закръглен до втория знак след десетичната запетая във формата: "Total income: {оборот за деня}"**.**

### Вход

* Ще получавате редове, докато не получите "end of shift".

### Изход

* За всеки валиден ред трябва да от печатате на конзолата:

"{име на клиент}: {продукт} - {оборот}"

* Когато получите "end of shift", отпечатайте оборота за деня, закръглен до втория знак след десетичната запетая във формата: "Total income: {оборот за деня}"**.**

### **Примери**

|  |  |  |
| --- | --- | --- |
| **Вход** | **Изход** | **Коментари** |
| %George%<Croissant>|2|10.3$  %Peter%<Gum>|1|1.3$  %Maria%<Cola>|1|2.4$  end of shift | George: Croissant - 20.60  Peter: Gum - 1.30  Maria: Cola - 2.40  Total income: 24.30 | Всяка команда е правилна, затова пресмятаме цената на всеки продукт.  Накрая отпечатваме оборота за деня |
| %InvalidName%<Croissant>|2|10.3$  %Peter%<Gum>1.3$  %Maria%<Cola>|1|2.4  %Valid%<Valid>valid|10|valid20$  end of shift | Valid: Valid - 200.00  Total income: 200.00 | На първия ред имаме невалидно потребителско име, затова пропускаме реда. На следващият ред няма количество за продукта. На третият ред няма валидна цена. Само на четвърти ред всичко е правилно |

## \*Енигма

Войната е в разгара си, но вие сте млад падуан, който може да обърне хода на войната с вашите умения по програмиране. Вашата задача е да направите програма, която декрептира съобщение от ордена и да спасите хиляди хора.

Ще получите няколко съобщения, които са **криптирани** с легендарната машина - Енигма. Трябва да **декрептирате съобщенията** чрез следните правила:

За да декрептирате правилно съобщенията, трябва да **намерите броя на всички букви** - **без значение** дали са **малки** или **големи**. След това за всеки символ трябва да **премахните броя на буквите** от криптираното съобщение.

След криптирането, съобщенията трябва да съдържат име на **планетата**, **население**, **тип атака** (**"A"** за атака и **"D"** за унищожение) и **брой войници**.

Името на планетата трябва да стартира с **"@"** и да съдържа **символи от латинската азбука**.

Броя на населението трябва да започва с **":"** и **число**.

Типът атака трябва да бъде **заобграден с "!"** (удивителен знак).

**Броя на войската** трябва да започва с "->" и число.

Реда на съобщенията трябва да бъде: **име на планета** -> **население** -> **тип атака** -> **брой на войска**. Всяка част ще бъде разделена по всички символи освен **: '@', '-', '!', ':' and '>'**.

### Вход

* На първия ред ще получите числото **n** - броя съобщения.
* На следващите **n** реда ще получите криптираните съобщения.

### Изход

След декрептирането, трябва да отпечатате информацията във формата:

На първия ред трябва да отпечатате планетите с тип атака и след това с тип унищожаване.

"Attacked planets: {Броя на планетите с тип атака}"  
"-> {име на планета}"  
"Destroyed planets: {Броя на планетите с тип унищожение}"  
"-> {име на планета}"

Планетите трябва да бъдат **сортинари по имена**.

### Бележки

* Числото **n** ще бъде в обхвата **[1...10]**

### Примери

|  |  |  |
| --- | --- | --- |
| **Вход** | **Изход** | **Коментари** |
| 2  **ST**CDoghudd4=63333$D$0**A**53333  EHf**s**y**ts**nhf?8555&I&2C9555**SR** | Attacked planets: 1  -> Alderaa  Destroyed planets: 1  -> Cantonica | Получаваме само две съобщения. Първото съобщение има сума 3, затова от всяка буква изваждаме с 3.  **PQ@Alderaa1:30000!A!->20000**  Второто съобщение има сума 5.  **@Cantonica:3000!D!->4000NM**  И двете съобщения имат валидни планети, населения, тип атаки и войски.  След декрептирането отпечатваме всяка планета във формата по-горе. |
| **Вход** | **Изход** | **Коментари** |
| 3  **tt**(''DG**s**vywge**r**x>6444444444%H%1B9444  GQh**rr**|**A**977777(H(**TTTT**  EHf**s**y**ts**nhf?8555&I&2C9555**SR** | Attacked planets: 0  Destroyed planets: 2  -> Cantonica  -> Coruscant | Получаваме три съобщения. Първото съобщение има сума 4  **pp$##@Coruscant:2000000000!D!->5000**  Второто съобщение има сума 7.  **@Jakku:200000!A!MMMM**  Съобщението е невалидно, няма валидни войски.  Третото съобщение има сума 5  **@Cantonica:3000!D!->4000NM** |

## \*Долни царства

Големи битки предстоят. В долните царства демоните се бият в дуел като само един ще уцелее.

Възложено ви е да запишете всичките участници в книгата на демоните на на могъщата битка на долното царство. Трябва да сортирате имената в кнгигата.

В **името** на всеки демон се съдържа **неговия живот** и **сила**.

Сумата на **asci кодовете на всички букви** (с изключение на числата (**0-9**), аритметични символи (**'+', '-', '\*', '/'**) и точка(**'.'**)) дава **живота** на демона.

Сумата на неговите цифри в името му е неговата сила. Обърнете внимание, че също така трябва да вземете предвид и знаците плюс '+' и минус '-' (+10 е 10 и -10 е -10). Също така има и символите ('\*' и '/'), които могат допълнително да променят неговата **сила като я умножават или я разделят на 2** (примерно неговото име е  
"m**15**\*/c**-5.0**", неговата сила е 15 + (-5,0)=10 и по-късно умножаваме по 2 (10\*2=20) и разделяме на 2 (20/2=10) )).

В заключение умножаваме и делим само, когато цифрите са включени в пресмятанията в **реда както в името**.

Ще получите всички демони на **един ред**, разделени със запетая, нула или интервал. Сортирайте ги по **азбучен ред** и отпечатайте техните имена **със силата и живота им**.

### Вход

Входа ще се чете от конзолата. Ще бъде на **един ред**, които съдържа имента на демоните, **разделени със запетая, нули и интервали** във формата: **"{Име на демон1}, {Име на демон2}, … {Име на демон3}"**

### Изход

Отпечатайте имената на демоните сортирани по име, разделени на нов ред във формата:

* **"{Име на демон} - {живот} health, {сила} damage"**

### Бележки

* Името на демона ще съдържа поне **един символ**
* В името на демона **няма да има** интервали или запетай
* Числото винаги ще съдържа цифри преди и след десетичната запетая.
* **Числото** в името му винаги е цяло или плаващо (с точка). Например валидни числа са '4', '+4', '-4', '3.5', '+3.5', '-3.5'

### Примери

|  |  |  |  |
| --- | --- | --- | --- |
| **Вход** | **Изход** | **Коментари** | |
| M3ph-0.5s-0.5t0.0\*\* | M3ph-0.5s-0.5t0.0\*\* - 524 health, 8.00 damage | M3ph-0.5s-0.5t0.0\*\*:  Живот= 'M' + 'p' + 'h' + 's' + 't' = 524 живот.  Сила= (3 + (-0.5) + (-0.5) + 0.0) \* 2 \* 2 = 8 жила. | |
| **Вход** | **Изход** | | **Коментари** |
| M3ph1st0\*\*, Azazel | Azazel - 615 health, 0.00 damage  M3ph1st0\*\* - 524 health, 16.00 damage | | Azazel:  Живто - 'A' + 'z' + 'a' + 'z' + 'e' + 'l' = 615 живот. Сила- няма числа = 0 сила.  M3ph1st0\*\*:  Живот- 'M' + 'p' + 'h' + 's' + 't' = 524 живот.  Сила - (3 + 1 + 0) \* 2 \* 2 = 16 сила. |
| Gos/ho | Gos/ho - 512 health, 0.00 damage | |  |

## \*Извличане на имейли

Напишете програма, която извлича всички имейли от даден текст. Текстът ще бъде на един ред. Отпечатайте имейлите на конзолата на нов ред. Емейл се зачита само, когато е във формата   
<потребителско име >@<домейн>, и където:

* **<потребителско име>** е редица от **символи** и **числа**, където '.', '-' и '\_' може да се появяват.
  + Валидни потребителско име са: "**stephan**", "**mike03**", "**s.johnson**", "**st\_steward**", "**softuni-bulgaria**", "**12345**".
  + Невалидни потребителски имена са: ''**--123**", "**.....**", "**nakov\_-**", "**\_steve**", "**.info**".
* **<домейн>**  е редица от поне две думи, разделени с точка **'.'**. Всяка дума е редици от символи и може да има '**-**' между думите
  + Валидни **домейни са**: "**softuni.bg**", "**software-university.com**", "**intoprogramming.info**", "**mail.softuni.org**".
  + Невалидни **домейни са**: "**helloworld**", "**.unknown.soft.**", "**invalid-host-**", "**invalid-**".
* Валидни **имейли са**: **info@softuni-bulgaria.org**, **kiki@hotmail.co.uk**, **no-reply@github.com**, **s.peterson@mail.uu.net**, **info-bg@software-university.software.academy**.
* Невалидни **имейли са**: **--123@gmail.com**, **…@mail.bg**, **.info@info.info**, **\_steve@yahoo.cn**, **mike@helloworld**, **mike@.unknown.soft.**, **s.johnson@invalid-**.

### Примери

|  |  |
| --- | --- |
| **Вход** | **Изход** |
| Please contact us at: support@github.com. | support@github.com |
| Just send email to s.miller@mit.edu and j.hopking@york.ac.uk for more information. | s.miller@mit.edu  j.hopking@york.ac.uk |
| Many users @ SoftUni confuse email addresses. We @ Softuni.BG provide high-quality training @ home or @ class. –- steve.parker@softuni.de. | steve.parker@softuni.de |