



АЮСА

Software Testing Company

Автоматизация тестирования

Регулярные выражения

План лекции

- Понятие и применение
- Элементы языка регулярных выражений
 - Простейшее регулярное выражение
 - Метасимволы и их значение
 - . (Точка)
 - Символьные классы
 - Квантификаторы
 - | (Выбор)
 - Группы
- Литература

Понятие и применение

Проверка корректности пользовательского ввода

facebook [Log in to existing account](#)

Sign Up

It's free and always will be.

galina makhotkina

galinamakhotkina@gmailcom

galinamakhotkina@gmailcom

.....

Birthday

Month Day Year Why do I need to provide my birthday?

Female Male

By clicking Sign Up, you agree to our [Terms](#) and that you have read our [Data Use Policy](#), including our [Cookie Use](#).

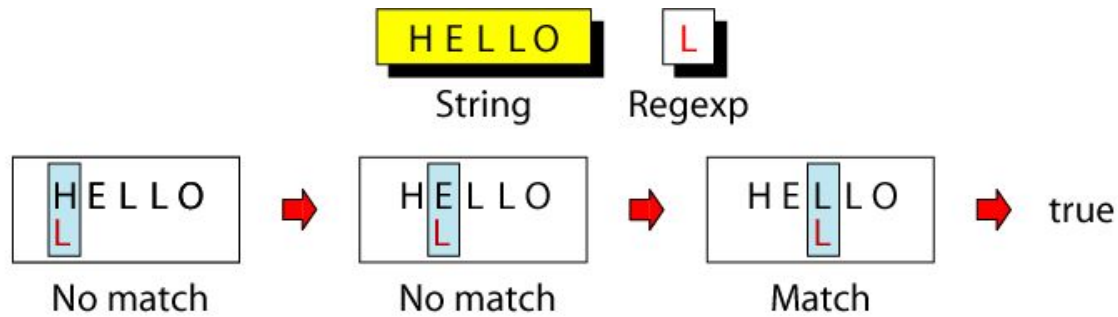
Sign Up

Please enter a valid email address.

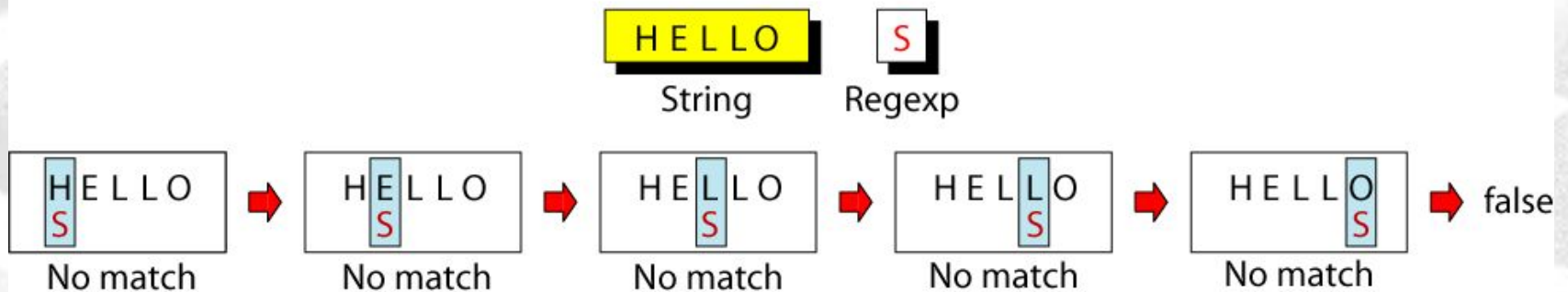
Mobile Find Friends Badges People Pages Places Apps Games Music
About Create Ad Create Page Developers Careers Privacy Cookies Terms Help

Facebook © 2013 · English (US)

Простейшее регулярное выражение



(a) Successful Pattern Match



(b) Unsuccessful Pattern Match

Метасимволы и их значение

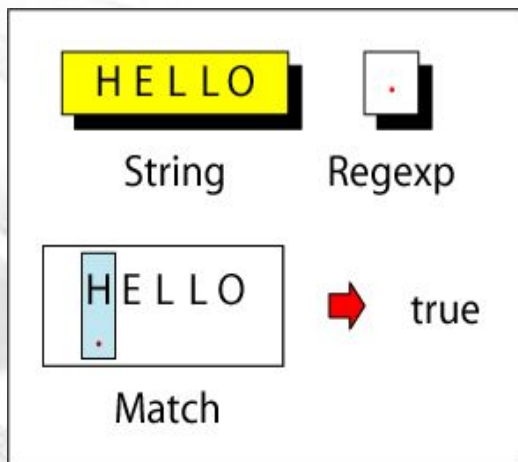
\wedge \$ \ . [] ? * + { } | ()



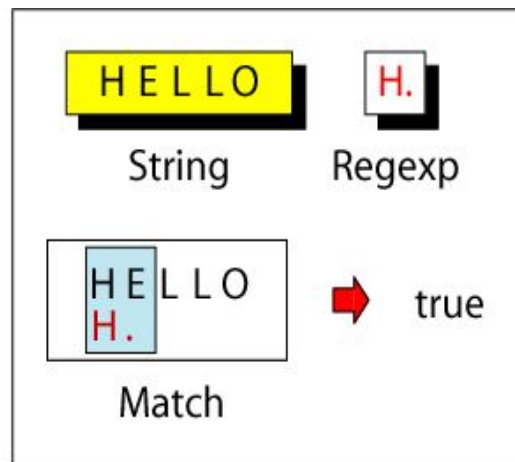
\. \^ \\$ \+ * и т. д.

. (Точка)

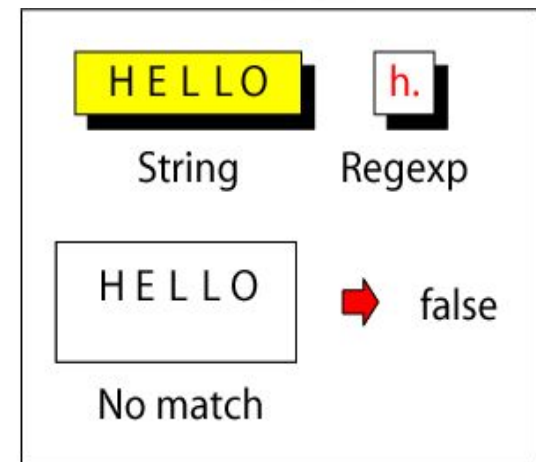
Определяет любой символ, кроме "\n" (новая строка).



(a) Single-Character



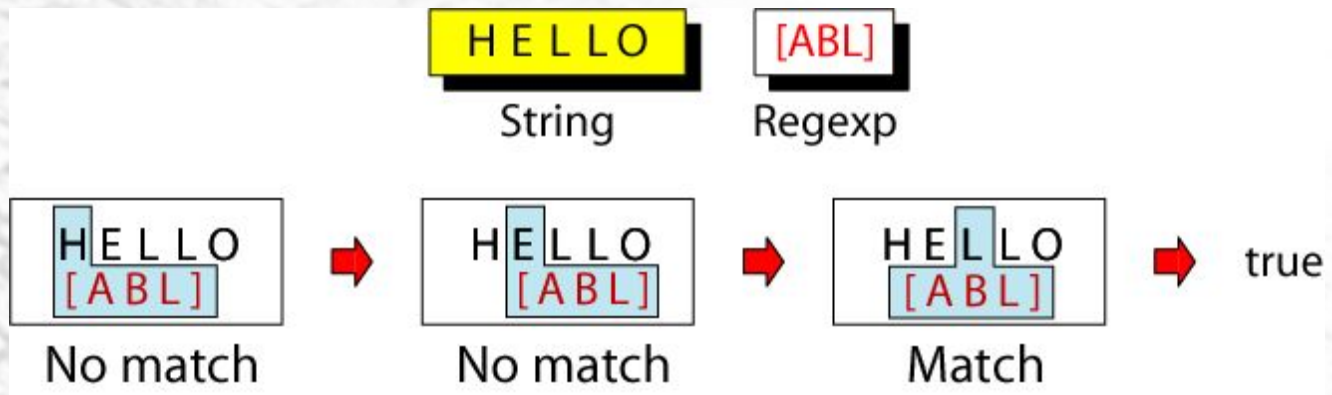
(b) Combination-True



(c) Combination-False

Символьные классы

- `[abcdef]` - Набор символов. Находит любой из указанных СИМВОЛОВ.



- `[^abcdef]` Исключающий набор символов. Находит любой из неуказанных символов.
- `[a-f]` Диапазон символов. Находит любой символ из диапазона.
- `[^a-f]` - Исключающий диапазон символов. Находит любой символ, не входящий в диапазон.

Символьные классы (сокращения)

.	любой символ
\d	любая цифра, т.е. [0-9]
\D	любой символ кроме цифры, т.е. [^\d]
\w	любой символ, образующий «слово», т.е. [A-Za-z0-9_]
\W	любой символ, НЕ образующий «слово», т.е. [^\w]
\s	любой whitespace символ, т.е. [.\t\v\r\n\f]
\S	любой НЕ whitespace символ, т.е. [^\s]

Можно использовать внутри символьного класса:
[\d\s] – цифры и whitespace символы

Квантификатор

Квантификатор	Описание	Пример
?	0 или 1 элемент {0,1}	colou?r
		color, colour colour
*	0 или более элементов {0,}	colou*r
		color, colour, colouur colour
+	1 или более элементов {1,}	colou+r
		colour, colouur color
{n}	ровно n элементов	co{2}l
		cool cl, col, coool
{m,n}	от m от n элементов	co{1,2}l
		col, cool cl, coool

| (Выбор)

gray | grey

gray, grey

other

gr (a | e) y

gray, grey

gry, graey

red | green | blue

red, green, blue

other

sw (i | a | u) m

swim, swam , swum

swiim, swiam...

Группы

Выделяет часть выражения в группу. Чтобы определить символы "(" и ")", следует использовать "\"(" и "\")".

A(BC){3}



ABCBCBC

(F(BC){2}G){2}



FBCBCGFBCBCG

Если в регулярном выражении необходимо сослаться на какую-то из предшествующих групп, следует использовать $\backslash n$, где вместо n подставить номер нужной группы.

regular expressions 101 — an online regex tester for javascript, php, pcre and python.

Your regular expression in: Python

" `([^\s]+(?:\.(jpg|gif|png))\.\2)` " flags: gmixsu

Your test string

`gffdgdgfg.jpg`

Your regular expression explained

- 1st Capturing group `([^\s]+(?:\.(jpg|gif|png))\.\2)`
 - `[^\s]+` match a single character not present in the list below
 - Quantifier: Between one and unlimited times, as many times as possible, giving back as needed [greedy]
 - `\s` match any white space character `[\r\n\t\f]`
 - `(?:\.(jpg|gif|png))` Positive Lookahead - Assert that the regex below can be matched
 - `\.` matches the character `.` literally
 - 2nd Capturing group `(jpg|gif|png)`
 - 1st Alternative: `jpg`
 - `jpg` matches the characters `jpg` literally (case sensitive)
 - 2nd Alternative: `gif`
 - `gif` matches the characters `gif` literally (case sensitive)
 - 3rd Alternative: `png`
 - `png` matches the characters `png` literally (case sensitive)
 - `\.` matches the character `.` literally
 - `\2` matches the same text as most recently matched 2nd backreference

Your regular expression explained

- 1st Capturing group `([^\s]+(?:\.(jpg|gif|png))\.\2)`
 - `[^\s]+` match a single character not present in the list below
 - Quantifier: Between one and unlimited times, as many times as possible, giving back as needed [greedy]
 - `\s` match any white space character `[\r\n\t\f]`
 - `(?:\.(jpg|gif|png))` Positive Lookahead - Assert that the regex below can be matched
 - `\.` matches the character `.` literally
 - 2nd Capturing group `(jpg|gif|png)`
 - 1st Alternative: `jpg`
 - `jpg` matches the characters `jpg` literally (case sensitive)
 - 2nd Alternative: `gif`
 - `gif` matches the characters `gif` literally (case sensitive)
 - 3rd Alternative: `png`
 - `png` matches the characters `png` literally (case sensitive)
 - `\.` matches the character `.` literally
 - `\2` matches the same text as most recently matched by the 2nd backreference

Match groups

MATCH 1		
1.	[0-12]	<code>gffdgdgfg.jpg</code>
2.	[9-12]	<code>jpg</code>

http://gskinner.com/RegExr/

RegExr - Mozilla Firefox

Файл Правка Вид Журнал Закладки Инструменты Справка

RegExr

gskinner.com/RegExr/

Match Replace

`\be(\w*)s\b`

global ignoreCase extended dotall multiline [Share Link](#)

Welcome to RegExr 0.3b, an intuitive tool for learning, writing, and testing Regular Expressions. Key features include:

- * real time results: shows results as you type
- * code hinting: roll over your expression to see info on specific [elements](#)
- * detailed results: roll over a match to see details & view group info below
- * built in regex guide: double click [entries](#) to insert them into your expression
- * online & desktop: [regexpr.com](#) or download the desktop version for Mac, Windows, or Linux
- * save your [expressions](#): My Saved [expressions](#) are saved locally
- * search Community [expressions](#) and add your own
- * create Share Links to send your [expressions](#) to co-workers or link to them on Twitter or your blog [ex. <http://RegExr.com?2rj16>]

Built by [gskinner.com](#) with Flex 3 [adobe.com/go/flex] and Spelling Plus Library for text highlighting [gskinner.com/products/sp1].

RegEx: `/\be(\w*)s\b/g`
pattern: `\be(\w*)s\b`
flags: `g`
1 capturing groups:
group 1: `(\w*)`

0.3.1b built by [gskinner.com](#) [About RegExr](#) [Desktop Version](#)

Samples My Saved Community


show all

- .
- \w
- \W
- \d
- \D
- \s
- \S
- [ABC]
- [^ABC]

Matches any character, except for line breaks if dotall is false.

*Некоторые люди, сталкиваясь с проблемой, думают:
"Знаю, я воспользуюсь регулярными выражениями".
Теперь у них две проблемы.*

Джейми Завински

Благодарю за
внимание  Software Testing Company