

Regular expressions cheat sheet

Basisregels

1. De *eerste* match telt
2. Matching is *case sensitive*
3. Standaard quantifiers zijn *greedy*

Metacharacters

De volgende tekens moet je escaper met een backslash '\' als je ze letterlijk wil matchen:

^ \$ () [\ | . * + ?

Alternation

notatie	betekenis
een twee	matcht een of twee

Quantifiers

betekenis	greedy	lazy
nul of meer	*	*?
één of meer	+	+?
nul of één	?	??
precies n	{n}	n.v.t.
minstens n	{n,}	{n,}?
hoogstens n	{,n}	{,n}?
n tot en met m	{n,m}	{n,m}?

Greedy: match *zoveel* mogelijk tekens.

Lazy: match zo *min* mogelijk tekens.

Grouping

notatie	betekenis
(r)	r is een groep; wat matcht wordt onthouden (<i>capturing group</i>)
(?:r)	r is een groep; wat matcht wordt niet onthouden (<i>non-capturing group</i>), en niet genummerd

Character classes

notatie	betekenis
\s	whitespace (space, tab, EOLs;\n en \r, en formfeed)
\S	niet whitespace
\w	woordcharacter (letter, cijfer, underscore)
\W	niet een woordcharacter
\d	cijfer
\D	niet een cijfer
[. . .]	verzameling <i>gewenste</i> tekens en/of combinatie van character classes
[^ . . .]	verzameling <i>ongewenste</i> tekens en/of combinatie van character classes
[a-z]	range van a tot en met z

Anchors

notatie	betekenis
^	begin van de regel
\$	eind van de regel
\b	woordgrens
\B	niet een woordgrens

Backreferences

Telling: capturing groups zijn genummerd vanaf 1 in de volgorde van de haakjes-openen.

notatie	betekenis
\k	gematchte tekst van de k-de groep (alleen binnen regex), met $1 \leq k \leq 9$
\$k	gematchte tekst van de k-de groep (alleen buiten regex).

Flags

Flags plaatsen achter de regex: `/regex/flags`

flag	betekenis	uitleg
i	case insensitive	geen onderscheid hoofd- en kleine letters
g	global search	doorgaan na eerste match (bv. bij <i>search & replace</i>)
s	single-line text	. matcht ook \n
m	multi-line anchors	^ matcht ook achter \n \$ matcht ook vóór \n
x	comment mode	white space negeren; commentaar mag achter #

Backtracking

1. Probeer te matchen vanaf begin.
2. Meerdere mogelijkheden? Onthoud huidige plek en mogelijkheden.
3. Probeer er één.
4. Geen match? Keer terug naar laatst onthouden plek en probeer volgende mogelijkheid.
5. Geen mogelijkheden meer? Ga verder met matchen totdat ofwel een match is gevonden, ofwel de hele match faalt.

Lookaround assertions

notatie	betekenis	naam
(?=r)	rechts hiervan moet <i>r</i> matchen	positive lookahead
(?!r)	rechts hiervan mag <i>r</i> niet matchen	negative lookahead
(?<=r)	links hiervan moet <i>r</i> matchen	positive lookbehind
(?<!r)	links hiervan mag <i>r</i> niet matchen	negative lookbehind