

Reguläre Ausdrücke: Sonderzeichen und Optionen

Hilfe: `perldoc perlre`

Sonderzeichen der Mustersuche

Die nächste Tabelle enthält Sonderzeichen in der Mustersuche. Wenn die Zeichen literal interpretiert werden sollen, müssen sie mit Backslash geschützt werden.

Sonderzeichen	Verwendung
\	schützt das nächste Zeichen
	trennt Alternativen
.	Wildcard, ausser Newlinezeichen (unter Option <code>s</code> einschließlich <code>\n</code>)
^	Anfang eines Strings (unter Option <code>m</code> : Anfang einer Zeile)
\$	Ende eines Strings (unter Option <code>m</code> : Ende einer Zeile)
*	Quantifikator: so oft es geht oder minimal 0 mal
+	Quantifikator: so oft es geht oder minimal 1 mal
?	Quantifikator: 1 mal oder minimal 0 mal
{n}	Quantifikator: genau <i>n</i> mal
{n, }	Quantifikator: so oft es geht, mindestens <i>n</i> mal
{n, m}	Quantifikator: <i>m</i> mal oder bis minimal <i>n</i> mal
[]	Zeichenklasse (Sonderzeichen innerhalb v. Zeichenklasse s.u.)
()	Gruppierung, Speicherklammer
(?:)	nur Gruppierung, keine Speicherklammer

Zusätzlich zu den im Kurs behandelten *gierigen* Quantifikatoren oben, die möglichst viele Zeichen matchen, gibt es auch die minimalen Varianten, die die minimale Erfüllung des Suchmusters bevorzugen.

Sonderzeichen	Verwendung
*?	Quantifikator: 0 oder mehrmals
+	Quantifikator: 1 oder mehrmals
??	Quantifikator: 0 oder 1 mal
{n, }?	Quantifikator: mindestens <i>n</i> mal
{n, m}?	Quantifikator: <i>n</i> mal, bis zu maximal <i>m</i> mal

Die nächste Tabelle enthält die besonderen Escape-Zeichen, die in den regulären

Ausdrücken **zusätzlich** zu den normalen String-Escape-Zeichen (`\t`, `\n`, Backslash) verwendet werden können.

Escape-Zeichen	Verwendung
<code>\A</code>	Anfang eines Strings
<code>\Z</code>	Ende eines Strings
<code>\w</code>	Wortzeichen
<code>\W</code>	Nichtwortzeichen
<code>\b</code>	Wortgrenze
<code>\B</code>	Nichtwortgrenze
<code>\s</code>	Whitespace-Zeichen
<code>\S</code>	Nicht-Whitespace-Zeichen
<code>\d</code>	Ziffer
<code>\D</code>	Nichtziffer
<code>\1</code>	Rückwärtsreferenz auf 1. Speicherklammer
<code>\2</code>	Rückwärtsreferenz auf 2. Speicherklammer
<code>\n</code>	Rückwärtsreferenz auf n. Speicherklammer

Nicht im Kurs behandelt haben wir die Escapes für Oktal- und Hexadezimalzahlen.

Escape-Zeichen	Verwendung
<code>\0nn</code>	Oktalzahl Format: Null, gefolgt von zwei beliebigen Ziffern
<code>\xnn</code>	Hexadezimalzahl Format: x, gefolgt von zwei beliebigen Ziffern

Sonderzeichen in Zeichenklassen

Escape-Zeichen für Zeichenklassen, wie z.B. `\w` oder `\W`, sowie die Escape-Zeichen für Oktal- und Hexadezimal können auch innerhalb von komplexen Klassen gelistet werden. Alle anderen Zeichen verlieren ihren Sonderstatus bis auf folgende Ausnahmen.

Sonderzeichen	Verwendung
<code>\</code>	schützt das nächste Zeichen
<code>]</code>	markiert Ende der Zeichenklasse
<code>-</code>	Bereichsoperator (nur im Kontext zwischen anderen Zeichen)
<code>^</code>	Negation (weiter Skopus über alle Zeichen der Klasse; nur im Kontext direkt hinter der öffnenden Klammer)

Modifikatoren der Mustersuche

- Syntax
Nachgestellte Option nach Suchausdruck (oder Substitution)
`m/ regex /option;`

Option	Verwendung
g	sucht alle Vorkommnisse (nicht nur eines) g lobal search
i	sucht nach Groß- und Kleinbuchstaben case- i nsensitive search
m	mehrzeiliger String wird als mehrere Zeilen interpretiert: ^ und \$ finden internes \n m ultiple lines
o	kompiliert das Muster nur einmal (steigert die Effizienz) compile pattern o nce
s	potentiell mehrzeiliger String wird als eine Zeile interpretiert: Wildcard (.) schließt \n mit ein string as s ingle line
x	erweitert die regulären Ausdrücke: kann Kommentare und Whitespace-Zeichen mit einschließen e xtended regular expression
e	bewertet Ersetzung als Perl-Ausdruck (nur Substitution mit <code>s///</code>) e valuate right sides as expression

Für weitere, im Kurs nicht behandelte Sonderzeichen und Optionen siehe die einschlägige Literatur.

Quellen

- Pokorra, Gerd, 2001. Perl. Eine Einführung. Hannover, RRZN-Handbuch. Kapitel 8
- Schamberger, Bernd. Perlguide (www.perlboard.de/perlguide/Inhalt.html)
- Larry Wall, Tom Christiansen & Jon Orwant, 2000. *Programming perl*, O'Reilly & Associates, Inc., 3rd edition. Kapitel 2