Ahmed Abo Bakr ELdesoky Salem
4211041
B2

A **regular expression** (shortened as **regex** or **regexp**; also referred to as **rational expression**) is a sequence of characters that specifies a *search pattern* in text. Usually such patterns are used by string-searching algorithms for "find" or "find and replace" operations on strings, or for input validation. It is a technique developed in theoretical computer science and formal language theory.

The concept of regular expressions began in the 1950s, when the American mathematician Stephen Cole Kleene formalized the description of a *regular language*. They came into common use with Unix text-processing utilities. Different syntaxes for writing regular expressions have existed since the 1980s, one being the POSIX standard and another, widely used, being the Perl syntax.

The phrase *regular expressions*, or *regexes*, is often used to mean the specific, standard textual syntax for representing patterns for matching text, as distinct from the mathematical notation described below. Each character in a regular expression (that is, each character in the string describing its pattern) is either a metacharacter, having a special meaning, or a regular character that has a literal meaning. For example, in the regex `b.`, 'b' is a literal character that matches just 'b', while '.' is a metacharacter that matches every character except a newline. Therefore, this regex matches, for example, 'b%', or 'bx', or 'b5'. Together, metacharacters and literal characters can be used to identify text of a given pattern or process a number of instances of it. Pattern matches may vary from a precise equality to a very general similarity, as controlled by the metacharacters. For example, `.` is a very general pattern, `[a-z]` (match all lower case letters from 'a' to 'z') is less general and `b` is a precise pattern (matches just 'b'). The metacharacter syntax is designed specifically to represent prescribed targets in a concise and flexible way to direct the automation of text processing of a variety of input data, in a form easy to type using a standard ASCII keyboard.

Regexes are useful in a wide variety of text processing tasks, and more generally string processing, where the data need not be textual. Common applications include data validation, data scraping (especially web scraping), data wrangling, simple parsing, the production of syntax highlighting systems, and many other tasks.