Hello! RegEx

A regular expression (`regex`) is a way to describe a particular pattern of characters that a *regular expression engine* would attempt to match in a given input text.

Mathematician Stephen Cole Kleene invented (1956) regular languages using his mathematical notation called 'regular sets' which then entered into popular use from 1968 in two use cases: *pattern matching* in a text editor and lexical analysis in a compiler. Among the first appearances of regular expressions in a programing form was seen when *Ken Thompson* built Kleene's notation into the editor *QED* as a means to match patterns in text files. In the 1980s the more complicated regexes started to arise in *Perl* and then in 1997, *Philip Hazel* developed *Perl Compatible Regular Expressions*, which attempts to closely mimic Perl's regex functionality and is currently used by many modern tools including PHP and Apache HTTP Server. Today regexes are widely supported in programming languages, text processing programs (particular lexers), advanced text editors, and some other programs. This tutorial will give you **just enough** knowledge to read and understand this book, to be a master on RegEx, you need to explore relevant literature referenced at end of this book.

With regular expressions we can:

- **Search for particular items**, within a large body of text.
- **Replace particular items**
- **Validate input**, for example, a password meets certain criteria such as, a mix of uppercase and lowercase, digits and punctuation, etc. and

- **Coordinate actions**, for example process certain files in a directory, but only if they meet particular conditions.

In short Regexes are super useful in text processing tasks, and also in 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.

REGEX Types #

The IEEE POSIX standard has three sets of compliance: **Basic Regular Expressions (BRE)**, **Extended Regular Expressions (ERE)**, and **Simple Regular Expressions (SRE)**. The SRE is deprecated.