Regular expression matching is an essential tool in string manipulating programs and plays crucial roles in scripting languages. Even though the theory of regular languages is well established, the semantics of regular expression matching in programming languages is quite involved due to ambiguity of regular expressions and various exotic features.

In this talk, we focus on regular expression matching based on the strategy of Perl and formulate its semantics as a nondeterministic parser using the list monad. Furthermore, we show that the standard implementation of matching can be obtained just by replacing the monad with one for backtracking. Finally, we discuss the semantics of POSIX-compliant regular expression matching. The talk is partly based on [1].

References