The Scanner

- Reading only **one character** at a time the scanner transforms a sequence of characters into a sequence of symbols that has **no structure**.

- There are single-character symbols (`'<'`, `x`) and multi-character symbols (`'<='`,`'variable'`). All symbols are uniquely identified by a token (integer).

- The whole sequence of symbols is not stored. Only the last read symbol is remembered.

*The scanner performs a **membership test** on a sequence of characters. It checks whether the symbols it sees are **valid symbols** of the programming language.*

What are valid symbols?

- **What is needed:**
  A specification of valid symbols -> regular expression and a formalism to write such a specification.

- **In selfie…**
  the chosen formalism is EBNF (see in the grammar.md file).

The Scanner

*The scanner performs a **membership test** on a sequence of characters. It checks whether the symbols it sees are **valid symbols** of the programming language.*

- Reading only **one character** at a time the scanner transforms a sequence of characters into a sequence of symbols that has **no structure**.

- There are single-character symbols (`'<'`, `x`) and multi-character symbols (`'<='`,`'variable'`). All symbols are uniquely identified by a token (integer).

- The whole sequence of symbols is not stored. Only the last read symbol is remembered.