

How Can We Define Languages in Computer Science?

Several scientific methods have been developed to precisely define the syntactic rules of languages.

Backus-Naur Form (BNF):

A **meta-language** used to describe the syntax of languages.

<name>	meta-symbol
::=	definition
	alternative
\{expression\}_{min}^{max}	repetition (minimum and maximum repetitions can be specified using subscripts)
GOTO	terminal symbol (in quotes for clarity, can also use apostrophes instead)

1. Example: Syntax of License Plates

Let's start with a few typical examples and try to generalize:

ABC-935, GHT-234, HSD-333, AI-BB-654

Syntax definition: $\$ \$ \langle \text{license_plate} \rangle ::= \langle \text{newType} \rangle \mid \langle \text{oldType} \rangle \mid \langle \text{oldType} \rangle ::= \{ \langle \text{letter} \rangle \}_3^3 - \{ \langle \text{number} \rangle \}_3^3 \mid \langle \text{newType} \rangle ::= \{ \langle \text{letter} \rangle \}_2^2 - \{ \langle \text{letter} \rangle \}_2^2 - \{ \langle \text{number} \rangle \}_3^3$
 $\mid \langle \text{letter} \rangle ::= A|B|C\dots|Z \mid \langle \text{number} \rangle ::= 0|1|2|3|4|5|6|7|8|9 \mid \$ \$$

2. Example: Syntax of Phone Calls in Hungary

Let's list a few examples and try to generalize:

062012345, 0036701234567, +36301234567, 0680460046

Syntax definition: $\$ \$ \langle \text{phone call} \rangle ::= \{ \langle \text{prefix} \rangle \}_0^1 \langle \text{city} \rangle \langle \text{customer} \rangle \mid \langle \text{prefix} \rangle ::= \{ +|00 \}_0^1 36|06 \mid \langle \text{city} \rangle ::= \{ \langle \text{number} \rangle \}_1^2 \mid \langle \text{customer} \rangle ::= \{ \langle \text{number} \rangle \}_6^7 \mid \langle \text{number} \rangle ::= 0|1|2|3|4|5|6|7|8|9 \mid \$ \$$

3. Example: How can we describe the BNF formula using itself?

$\langle \text{BN formula} \rangle ::= \langle \text{rule} \rangle \mid \langle \text{rule} \rangle \mid \langle \text{rule} \rangle ::= \langle \text{identifier} \rangle \langle \text{expression} \rangle \mid \langle \text{identifier} \rangle ::= \langle \text{letter} \rangle \mid \langle \text{letter} \rangle \mid \langle \text{digit} \rangle \mid \langle \text{expression} \rangle ::= \langle \text{term} \rangle \mid \langle \text{term} \rangle \mid \langle \text{term} \rangle ::= \langle \text{factor} \rangle \mid \langle \text{factor} \rangle \mid \langle \text{factor} \rangle ::= \langle \text{identifier} \rangle \mid \langle \text{terminal_symbol} \rangle \mid \langle \text{terminal_symbol} \rangle ::= \langle \text{terminal_symbol} \rangle$

$\backslash \text{angle} \backslash \text{text}\{\text{character}\} \backslash \text{range} \backslash \text{right}\rbrace \backslash \backslash [\backslash \text{angle} \backslash \text{text}\{\text{letter}\} \backslash \text{range} ::= \backslash \text{angle} \backslash \text{text}\{\text{uppercase}\} \backslash \text{range} \backslash \text{mid} \backslash \text{angle} \backslash \text{text}\{\text{lowercase}\} \backslash \text{range} \backslash \backslash [\backslash \text{angle} \backslash \text{text}\{\text{uppercase}\} \backslash \text{range} ::= A \backslash \text{mid} B \backslash \text{mid} C \backslash \text{dots} Z \backslash \backslash [\backslash \text{angle} \backslash \text{text}\{\text{lowercase}\} \backslash \text{range} ::= a \backslash \text{mid} b \backslash \text{mid} c \backslash \text{dots} z \backslash \backslash [\backslash \text{angle} \backslash \text{text}\{\text{digit}\} \backslash \text{range} ::= 0 \backslash \text{mid} 1 \backslash \text{mid} 2 \backslash \text{mid} 3 \backslash \text{mid} 4 \backslash \text{mid} 5 \backslash \text{mid} 6 \backslash \text{mid} 7 \backslash \text{mid} 8 \backslash \text{mid} 9 \backslash$

From:

<https://edu.iit.uni-miskolc.hu/> - **Institute of Information Science - University of Miskolc**

Permanent link:

https://edu.iit.uni-miskolc.hu/tanszek:oktatas:techcomm:bn_formulas

Last update: **2025/11/25 07:50**

