

# 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-1234, GHT-234, HSD-333, AI-BB-654

Syntax definition:  $\langle \text{license\_plate} \rangle ::= \langle \text{newType} \rangle \mid \langle \text{oldType} \rangle \mid \langle \text{oldType} \rangle \langle \text{newType} \rangle$   
 $\langle \text{newType} \rangle ::= \langle \text{letter} \rangle_3^3 - \langle \text{number} \rangle_3^3$   
 $\langle \text{oldType} \rangle ::= \langle \text{letter} \rangle_2^2 - \langle \text{letter} \rangle_2^2 - \langle \text{number} \rangle_3^3$   
 $\langle \text{letter} \rangle ::= A|B|C\dots|Z$   
 $\langle \text{number} \rangle ::= 0|1|2|3|4|5|6|7|8|9$

### 2. Example: Syntax of Phone Calls in Hungary

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

062012345, +36301234567, 0680460046

Syntax definition:  $\langle \text{phone call} \rangle ::= \langle \text{prefix} \rangle_0^1 \langle \text{city} \rangle \langle \text{customer} \rangle$   
 $\langle \text{prefix} \rangle ::= +\langle \text{number} \rangle_0^1$   
 $\langle \text{city} \rangle ::= \langle \text{number} \rangle_1^2$   
 $\langle \text{customer} \rangle ::= \langle \text{number} \rangle_6^7$   
 $\langle \text{number} \rangle ::= 0|1|2|3|4|5|6|7|8|9$

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

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

Last update: 2025/11/17 20:17 tanszek:oktatas:techcomm:bn\_formulas [https://edu.iit.uni-miskolc.hu/tanszek:oktatas:techcomm:bn\\_formulas?rev=1763410651](https://edu.iit.uni-miskolc.hu/tanszek:oktatas:techcomm:bn_formulas?rev=1763410651)

$\left\{ \begin{array}{l} \text{letter} \\ \text{uppercase} \\ \text{lowercase} \end{array} \right\} \text{ \rangle ::= A \mid B \mid C \dots Z }$   
 $\left\{ \begin{array}{l} \text{digit} \\ \text{lowercase} \end{array} \right\} \text{ \rangle ::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 }$

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?rev=1763410651](https://edu.iit.uni-miskolc.hu/tanszek:oktatas:techcomm:bn_formulas?rev=1763410651)

Last update: **2025/11/17 20:17**

