# Introduction to Context-Free Grammars

Ian Ludden

# Learning Objectives

By the end of this lesson, you will be able to:

# Learning Objectives

By the end of this lesson, you will be able to:

- Correctly interpret the definition of a context-free grammar: variables, rules, start symbols, terminal symbols.

# Learning Objectives

By the end of this lesson, you will be able to:

- Correctly interpret the definition of a context-free grammar: variables, rules, start symbols, terminal symbols.
- Correctly interpret a grammar rule, e.g., what does $\varepsilon$ or a vertical bar mean on the right hand side of a rule.

# Motivation, i.e., why you should care

# Motivation, i.e., why you should care

- Applications to natural language processing (NLP), a field with lots of recent applications of deep learning (Link to article)

# Motivation, i.e., why you should care

- Applications to natural language processing (NLP), a field with lots of recent applications of deep learning (Link to article)
- Context-free grammars will be covered in more depth in CS 374

# Motivation, i.e., why you should care

- Applications to natural language processing (NLP), a field with lots of recent applications of deep learning (Link to article)
- Context-free grammars will be covered in more depth in CS 374
- Good practice for inductive proofs on trees

# Formal Definition

# Formal Definition

## Definition

A ***context-free grammar*** $G = (\Sigma, \Gamma, R, S)$ is a structure defined by:

- A finite set $\Sigma$ of *symbols*, or ***terminals*** → "ending points"

# Formal Definition

## Definition

A ***context-free grammar*** $G = (\Sigma, \Gamma, R, S)$ is a structure defined by:

- A finite set $\Sigma$ of *symbols*, or ***terminals***
- A finite set $\Gamma$ of ***non-terminals*** (disjoint from $\Sigma$)

Variables

# Formal Definition

## Definition

A ***context-free grammar*** $G = (\Sigma, \Gamma, R, S)$ is a structure defined by:

- A finite set $\Sigma$ of *symbols*, or ***terminals***
- A finite set $\Gamma$ of ***non-terminals*** (disjoint from $\Sigma$)
- A finite set $R$ of ***production rules***

# Formal Definition

## Definition

A ***context-free grammar*** $G = (\Sigma, \Gamma, R, S)$ is a structure defined by:

- A finite set $\Sigma$ of *symbols*, or ***terminals***
- A finite set $\Gamma$ of ***non-terminals*** (disjoint from $\Sigma$)
- A finite set $R$ of ***production rules***
- A finite set $S$ of ***start symbols*** (often just one)

# Formal Definition

## Definition

A ***context-free grammar*** $G = (\Sigma, \Gamma, R, S)$ is a structure defined by:

- A finite set $\Sigma$ of *symbols*, or ***terminals***
- A finite set $\Gamma$ of ***non-terminals*** (disjoint from $\Sigma$)
- A finite set $R$ of ***production rules***
- A finite set $S$ of ***start symbols*** (often just one)

⟨sentence⟩ → ⟨noun phrase⟩⟨verb phrase⟩⟨noun phrase⟩
⟨noun phrase⟩ → ⟨adjective phrase⟩⟨noun⟩
⟨adj. phrase⟩ → ⟨article⟩ | ⟨possessive⟩ | ⟨adjective phrase⟩⟨adjective⟩
⟨verb phrase⟩ → ⟨verb⟩ | ⟨adverb⟩⟨verb phrase⟩

⟨noun⟩ → dog | trousers | daughter | nose | homework | time lord | pony | ···
⟨article⟩ → the | a | some | every | that | ···
⟨possessive⟩ → ⟨noun phrase⟩'s | my | your | his | her | ···
⟨adjective⟩ → friendly | furious | moist | green | severed | timey-wimey | little | ···
⟨verb⟩ → ate | found | wrote | killed | mangled | saved | invented | broke | ···
⟨adverb⟩ → squarely | incompetently | barely | sort of | awkwardly | totally | ···

# More Examples

## Example 1: Binary Strings

Start symbol is $S$, terminals are 0, 1, and $\varepsilon$, rules are:

$$S \rightarrow 0S$$
$$S \rightarrow 1S$$
$$S \rightarrow \varepsilon$$

Generate 010:

S
0S
01S
010S
010ε = 010

← empty string

$S \rightarrow 0S \mid 1S \mid \varepsilon$

## Example 2: Simple Arithmetic Expressions

Start symbols are $E$ and $V$ (also non-terminals), terminals are $x$, $y$, $+$, and $\times$, rules are:

E = expression
V = variable

$$E \rightarrow E + V \mid E \times V \mid V + V \mid V \times V$$
$$V \rightarrow x \mid y$$

x × y + y

E
① E + V
② V × V + y
③ x × y + y

# Recap: Learning Objectives

By the end of this lesson, you will be able to:

- Correctly interpret the definition of a context-free grammar: variables, rules, start symbols, terminal symbols.
- Correctly interpret a grammar rule, e.g., what does $\varepsilon$ or a vertical bar mean on the right hand side of a rule.