COMP4403 Crib Sheet

Parsing Theory

Context Free Grammars

Basic Example of Context Free Grammar
$E \to E\ Op\ E$
$E \to "("E")"$
$E \to number$
$Op \to "+"$
$Op \to "-"$
$Op \to "*"$
Has start symbol E, nonterminals $\{E, Op\}$, and terminals $\{\backslash(", \backslash)", number, \backslash+", \backslash-", \backslash*"\}$
A context-free grammar consists of:

- A finite set, $\sum$, of terminal symbols.
- A finite nonempty set of nonterminal symbols (disjoint from the terminal symbols).
- A finite nonempty set of producions of the form of $A \to \alpha$, where A is a nonterminal symbol, and α is a possibly empty sequence of symbols, each of which is either a terminal of nonterminal symbol.
- A start symbol that must be a nonterminal symbol

Directly Derives

If there is a production in the form of $N \to \gamma$ then we can directly derive $\alpha\ N\beta\ \to \alpha\ \gamma\ \beta$, where α and β are possibly empty sequences of terminal and nonterminal symbols.

Derives

Given a sequence of terminal and nonterminal symbols, α, derives a sequence β, written $\alpha \overset{*}{\Rightarrow} \beta$ if there is a finitie sequence of zero or more direct derivation steps that start from α and finishing with β, there must be one or more sequence $\gamma_1, \gamma_2, \ldots, \gamma_n$ such that $\alpha = \gamma_0 \Rightarrow \gamma_1 \Rightarrow \ldots \Rightarrow \gamma_n = \beta$. Note that zero steps are allowed.

Nullable

A possibly empty sequence of symbols, α, is nullable if $\alpha \overset{*}{\Rightarrow} \epsilon$ or $\alpha \overset{*}{\Rightarrow}$ Nullable rules

- $\epsilon is nullable$
- any terminal symbol is not nullable
- a sequence of symbols is nullable if all of its constructs are nullable
- a set of alternatives is nullable if any of its constructs are nullable
- EBNF constructs for optionals and repetitions are nullable
- a nonterminal is nullable if there is a production with a nullable right-hand side