Monday, August 23, 2021 11:47 AM

Programming language have

SYNTAX describe how the program "looks"

SEMANTICS describe the "meaning"

Quick introduction to Syntax

The cat drank the milk .

words (lexical units): "The" "cat" "drank" "the" "milk" "."

categories/kinds: DEFART NOUN VERB DEFART NOUN DOT

sentence components: SUBJECT VERB OBJECT DOT

SENTENCE

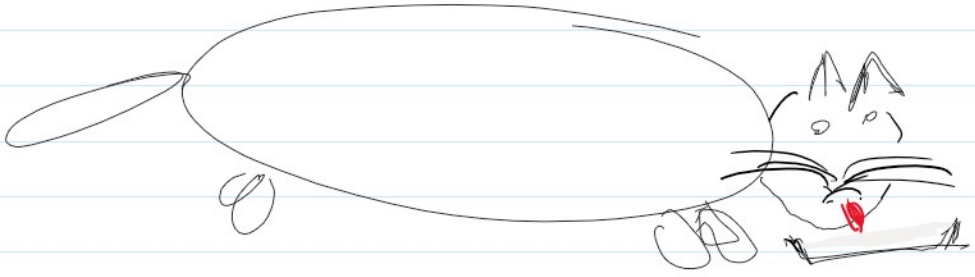
parsing:

1. identify words
2. identify kinds / categories of words

) lexical analysis

3. identify sentence components
4. identify sentence

) syntax analysis

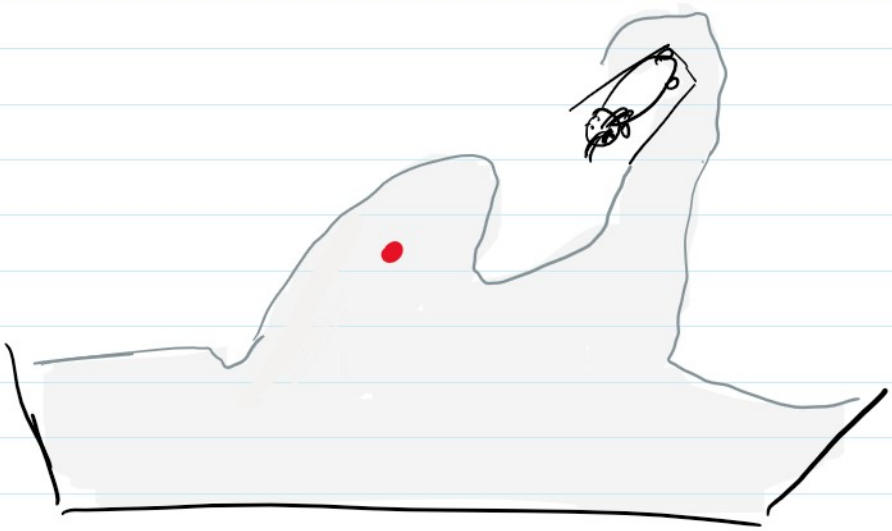


The milk drank the cat.

DEFART NOUN VERB DEFART NOUN DOT

SUBJECT VERB OBJECT DOT

SENTENCE



SYNTAX DOES NOT CARE ABOUT MEANING!

Syntax of programming languages

Alphabet a finite set of symbols

Example { a, b, c, !, & }

Example. Alphabet of the C language =

{ a, b, ..., z, A, ..., Z, 0, ..., 9,

{ a, b, - - , z, A, - - Z, 0, - - , 9,
! , c, >, =, - - }

String A string over an alphabet is a finite sequence of symbols from the alphabet. ↑ order matters

Example. abc! is a string over {a, b, c, d, &, !}
abcde is not a string over
{a, b, c, d, &, !}

Token A token (corresponds to categories or kinds of words) is a set of strings. We can think of the token as the name or label of the set.