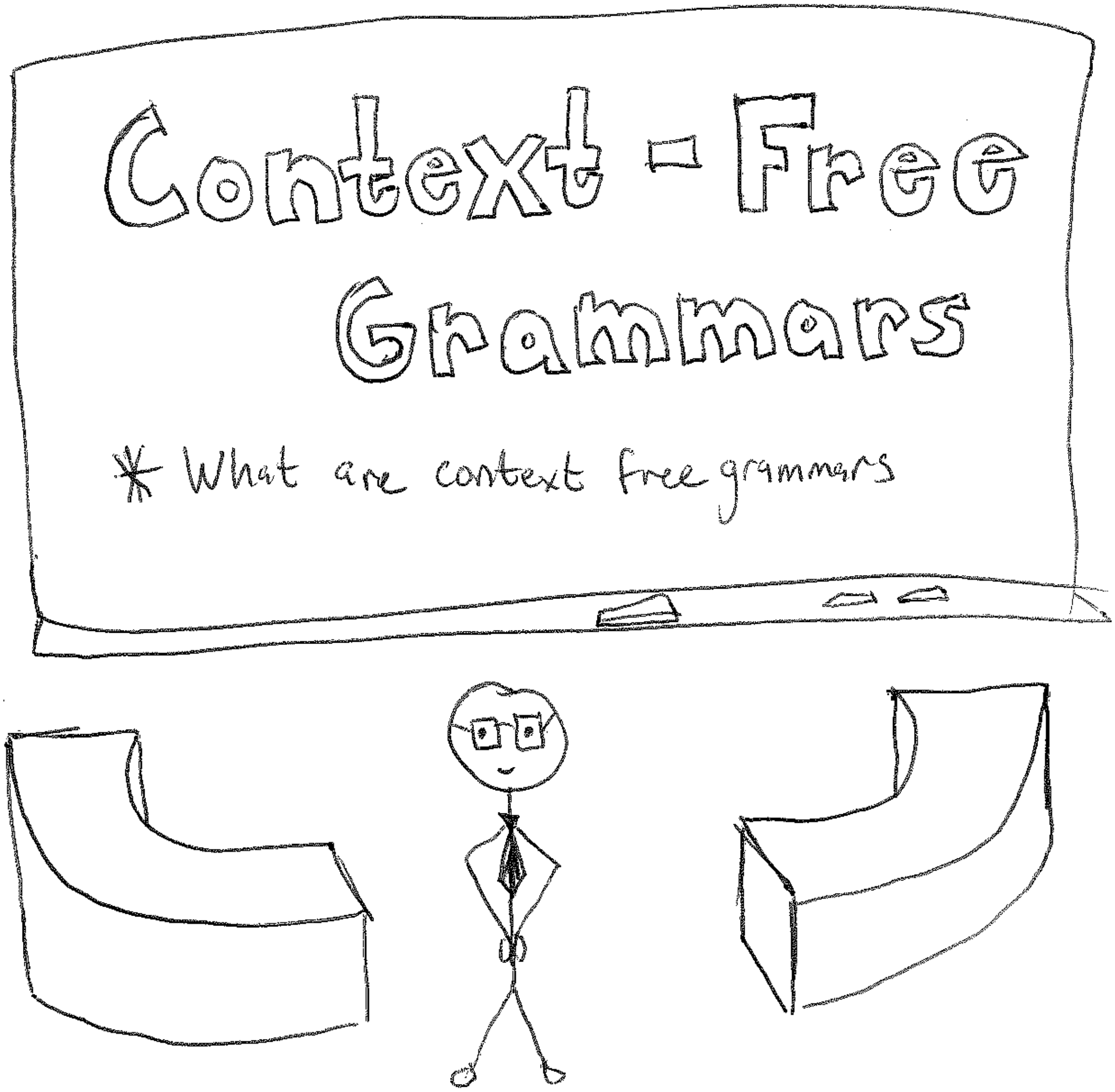
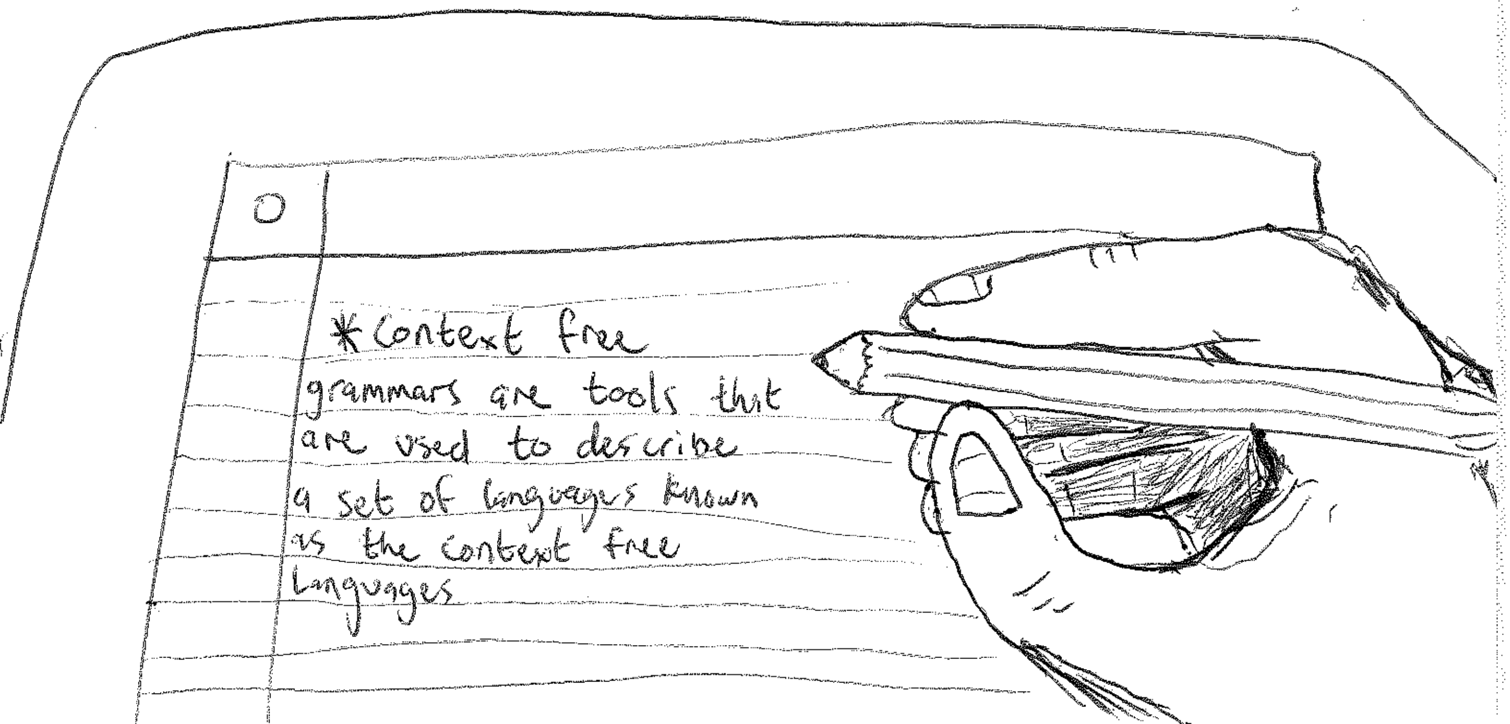
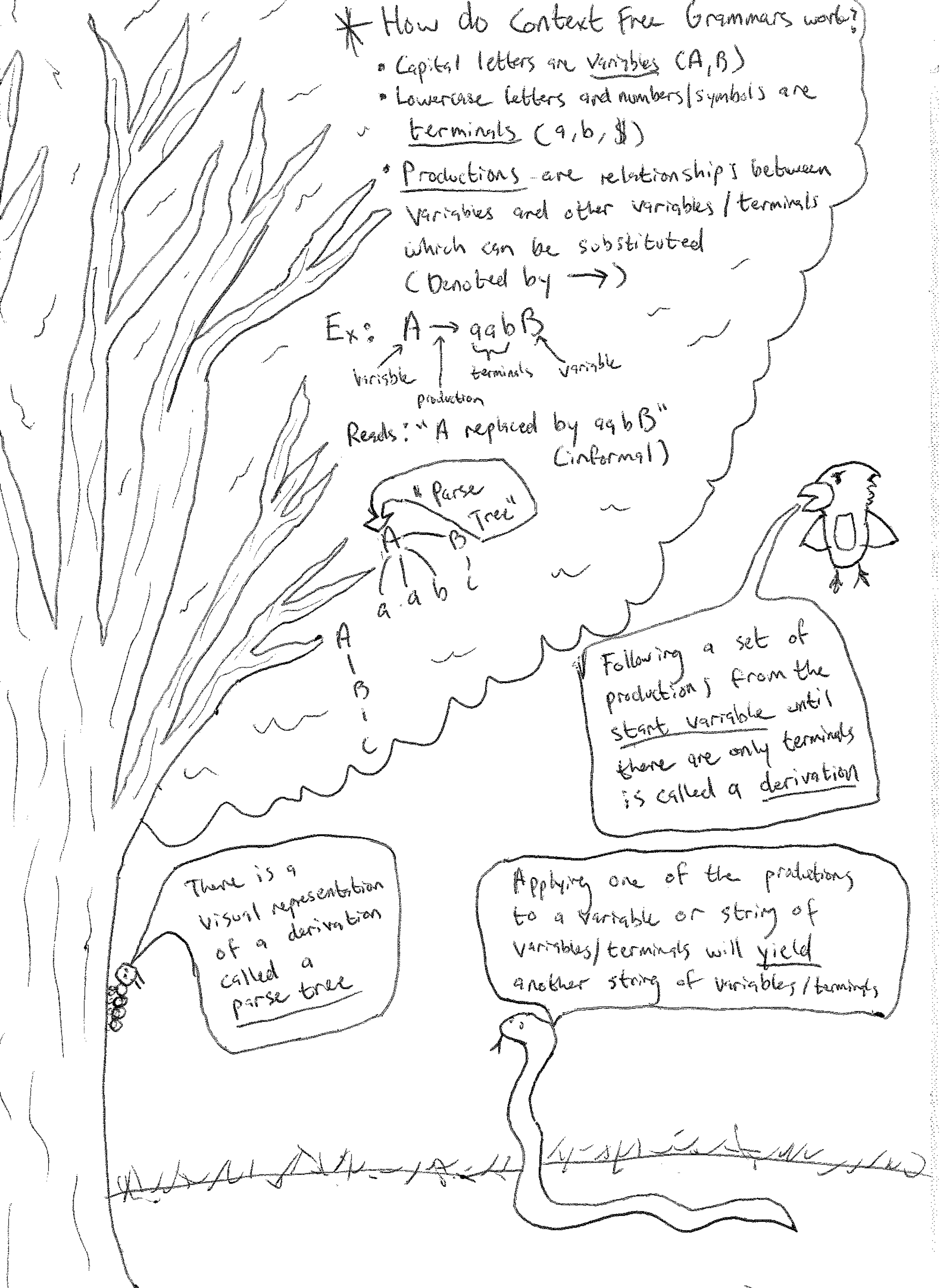

Context-Free Grammars
* What are context free grammars


* Context free grammars are tools that are used to describe a set of languages known as the context free languages



* How do Context Free Grammars work?

- Capital letters are <u>variables</u> (A, B)
- Lowercase letters and numbers/symbols are <u>terminals</u> (a, b, $)
- <u>Productions</u> are relationships between variables and other variables / terminals which can be substituted (Denoted by →)

Ex: $$\underset{\text{variable}}{A}\ \underset{\text{production}}{\rightarrow}\ \underbrace{aab}_{\text{terminals}}\underset{\text{variable}}{B}$$

Reads: "A replaced by aabB" (informal)

Following a set of productions from the <u>start variable</u> until there are only terminals is called a <u>derivation</u>

There is a visual representation of a derivation called a <u>parse tree</u>

Applying one of the productions to a variable or string of variables/terminals will <u>yield</u> another string of variables/terminals

Formal Definition

A context free grammar is a 4-tuple (V, Σ, R, S) where

1. V is a finite set called the variables
2. Σ is a finite set, disjoint from V, called the terminals
3. R is a finite set of rules, with each rule being a variable and a string of variables / terminals
4. $S \in V$ is the start variable

Example of a derivation

- $A \rightarrow Ba$
- $B \rightarrow Cab$
- $C \rightarrow c$

} Here are the set of rules which must be followed in the derivation

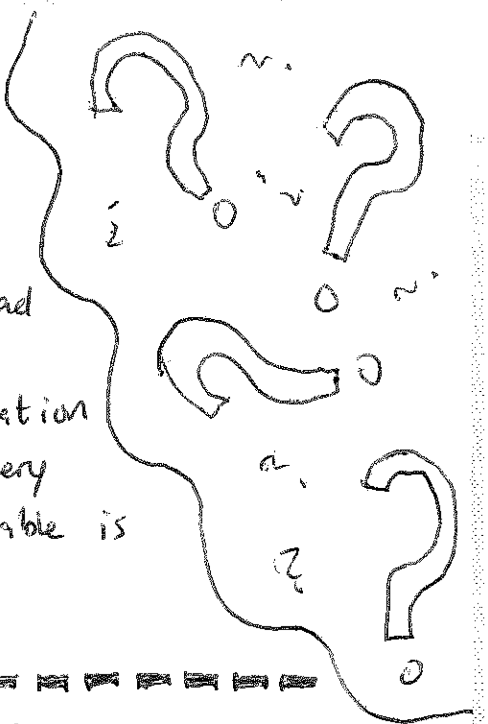
By convention, the start variable is A since it is the leftmost variable in the first rule

- Starting with A, $A \rightarrow Ba$ (A yields Ba, thus A is "replaced" with Ba)
- Now given Ba, one of the rules follows that $B \rightarrow Cab$,

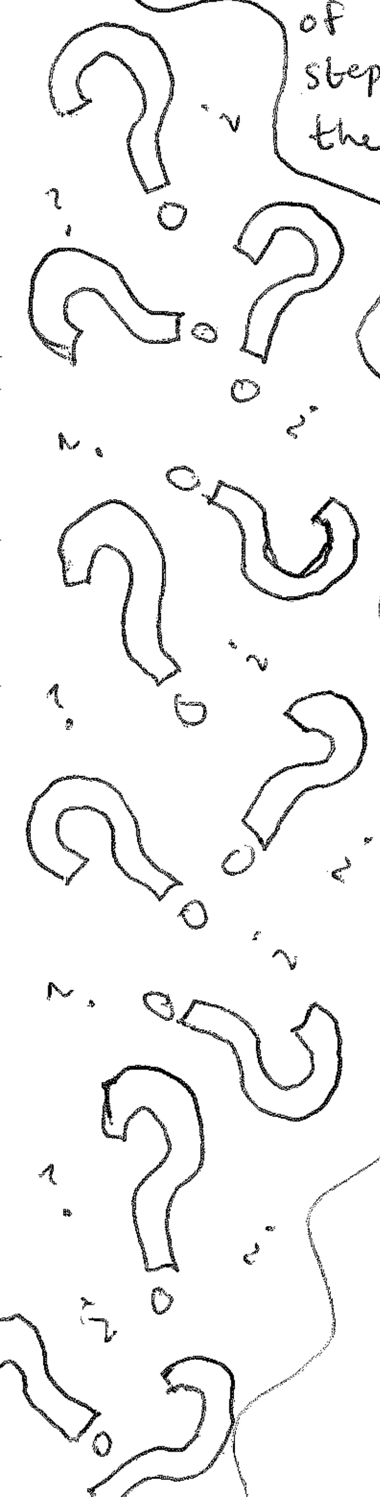
  thus: $Ba \rightarrow Caba$
- Given Caba, one of the rules states that $C \rightarrow c$,

  thus: $Caba \rightarrow caba$
- Since all that is left is terminals, we are finished and can claim that A derives caba.

* Ambiguity



- If a grammar can generate the same string in several ways, the grammar is called Ambiguous
- More formally, the grammar must contain at least two leftmost derivations which lead to the same string
  - A leftmost derivation is a derivation of a string w such that at every step, the leftmost remaining variable is the one replaced.



Example of determining Ambigoity

- Given the rules

A → (A) | (A)(A) | B
B → BB | AA | CC
C → AB | ε

In order to show that this grammar is ambiguous, we must show that at least two left most derivations lead to the same string

Derivation 1:

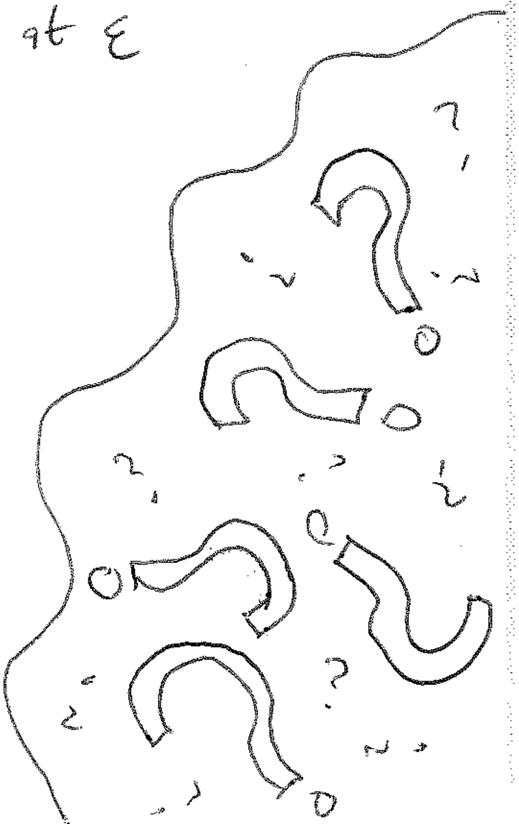
A → B
B → CC
CC → εC
C → ε

Arrives at ε

Derivation 2:

A → B
B → BB
BB → CCB
CCB → εCB
CB → εB
B → CC
CC → εC
C → ε

Arrives at ε



Since both leftmost derivations lead to the same string, this grammar is considered Ambiguous.