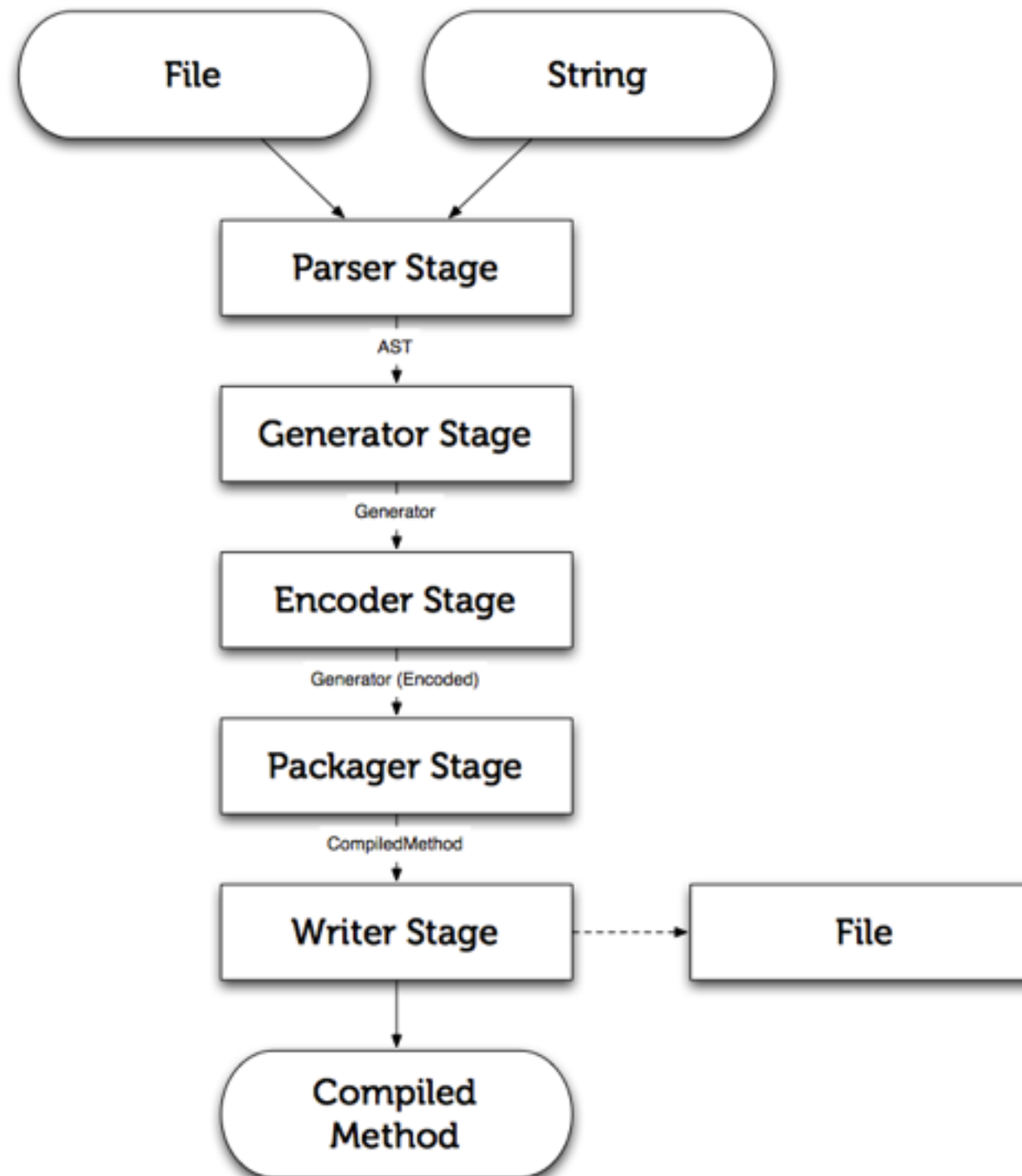


Lexical Analysis

(analysis of a simple programming
language called "natasha")

Anatomy of a compiler



What is Natasha?

- A simple programming language for doing basic arithmetic operations
- Written because I wanted to learn more about how compilers worked and also get a deeper understanding of how source code moves plain text to actual applications
- All "natasha" programs end with ".natasha"

Tokenization

; This is a comment

int a = 20

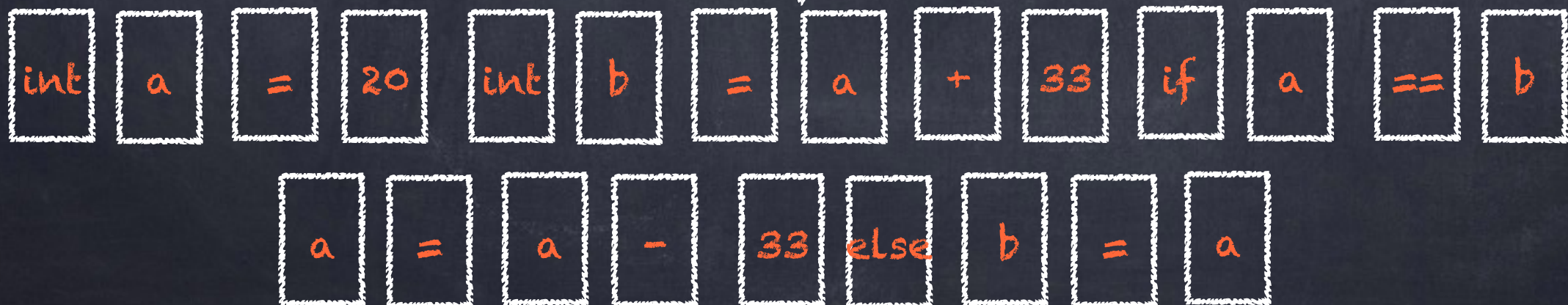
int b = a + 33

if a == b

 a = a - 33 ; this also another comment

else

 b = a



Details of the Lexer

- File extension is checked ".natasha"
- No processing is done if it is not a valid ".natasha" file
- heavily based on RegEx
- Unknown patterns causes compilation error

Demo Time