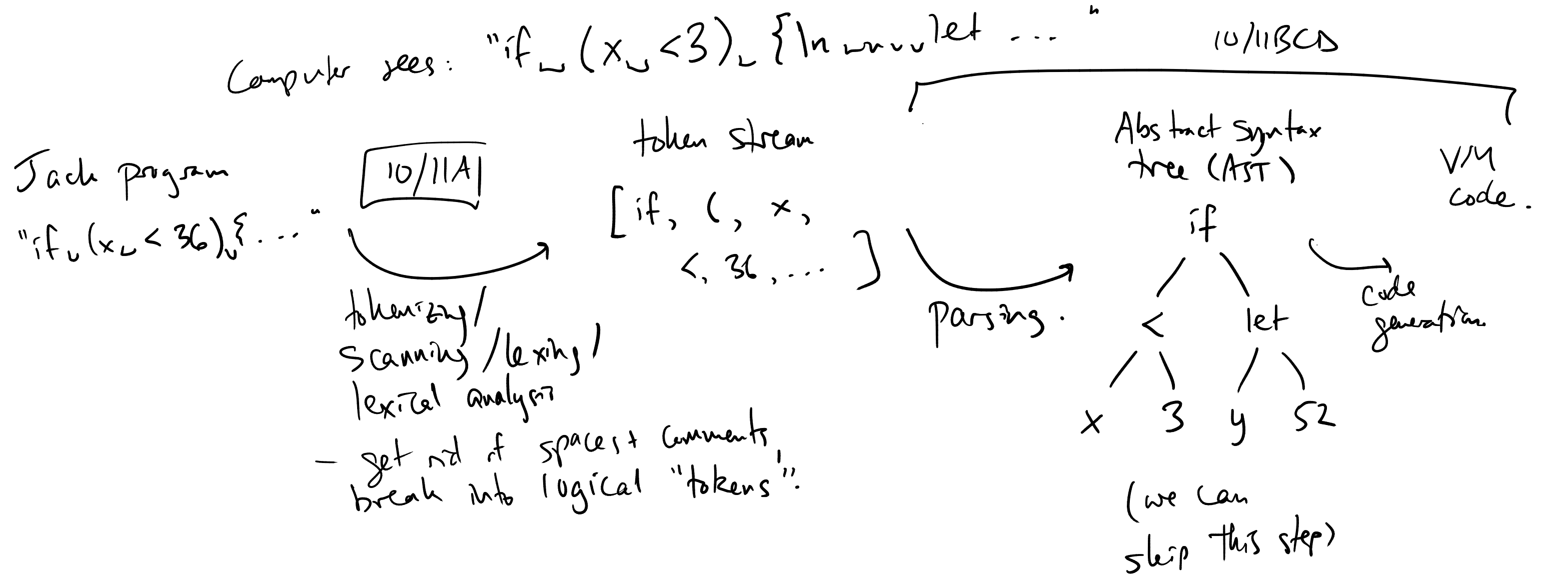
```
we see :      if (x < 3) {     // comment
                  let y= 52;
              }
```

Computer sees: "if␣(x␣<3)␣{\n␣␣␣␣let ... "

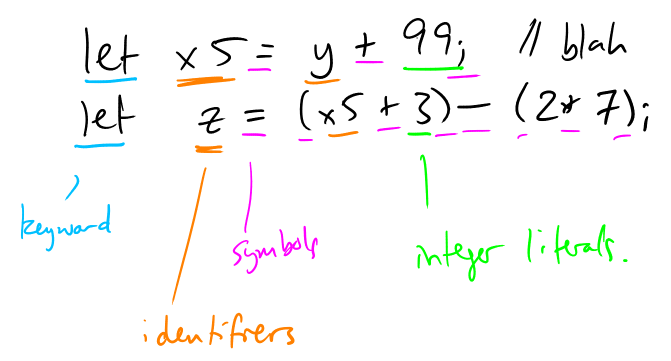


– get rid of spaces + comments, break into logical "tokens".

Tokenizing

Use a simple example language with:

- variables
- integers
- +, −, * , parens
- let statements
- comments.



Kinds of tokens:

Keywords (let)
Symbols (−, +, ;, =, etc.)
Identifiers (Names the programmer made up).
Integer literals (e.g. 99, 3, etc.)

Tokenizer

- init : take input text, initialize
- has_more_tokens: are there more tokens?
- advance: go to the next token.