

Prelude

Other uses of compiler technology mentioned by people in the class

- Used in CS410 to parse graphics model
- Used frequently throughout career to embed languages within applications
- domain-specific XML

People also noticed that...

- SableCC is also a compiler
- I wonder how it parses its input? Did it use a parser generator?



Plan for Today

How do lexer generators work?

- Convert regular expressions to NFA
- Converting an NFA to DFA

Specifying Tokens with SableCC

Theory meets practice:

- Regular expressions, formal languages, grammars, parsing...

SableCC example input file:

```
Package minijava;

Helpers
all = [0..0xFFFF];
cr = 13;

digit = ['0'..'9'];
letter = ['a'..'z'] | ['A'..'Z'];
underscore = '_';

not_star = [all - '*'];
not_star_slash = [not_star - '/'];

c_comment = '/*' not_star* ('*'
(not_star_slash not_star*)?)* '*/';

Tokens
plus = '+';
if = 'if';

id = letter (letter | digit | underscore)*;

blank = C' ' | eol | tab)*;
comment = c_comment | line_comment;

Ignored Tokens
blank,
comment;
```