**Reporting Errors Based on Syntax Tree**

Suppose we discover undclared variable ‘i’ in syntax tree

* in a program of 100K lines

What error message would you like compiler to give?

* An ocurrence of variable ‘i’ not declared
* An ocurrence of variable ‘i’ *in procedure P* not declared
* Variable ‘i’ undeclared at line 514, position 12

How would we emit this error message from syntax trees?

**One Pass Compiler**

One pass compiler (traditional, fast):

* parser does not construct trees
* instead, it directly generates code
* has symbol table storing sufficient information to generate code
* does only very simple optimizations

Lexer knows where undefined variable occurred - current position in lexer approximately indicates the position of the problem in source code

**Limitations of One-Pass Compilers**

No mutually recursive definitions

forward B(int x);

int A(int x) {

if (x < 10) return x;

else {

return B(x)+B(x-1);

}

}

int B(int x) {

if (x % 2 == 0)

return x;

else

return A(x)+1;

}

Either A or B unknown when reading linearly

* must give interface ahead of time to enable checks of use

More fundamentally: no interesting optimizations possible in one pass

**Modern Compilers**

Flow in modern compilers:

* parser builds syntax tree
* further processing done on syntax tree
* mutually recursive definitions not a problem

But users want error messages in terms of file position (line, column)

After tree is built, what is current position for lexer?

* end of file
* useless for error messages

Solution: store positions in syntax trees

Every node in concrete syntax tree has first and last token

Store positions inside abstract syntax tree during tree construction

**Syntax Trees with Positions**

Tree nodes store positions in file where they begin

Compiler uses these positions to report errors

For identifier nodes allows reporting variable uses

Variable 'i' in line 11, column 5 undeclared

For other nodes useful for type errors, e.g. could even report

(x + y) \* (!ok)

Type error in line 13,

expression in column 11-15 has type bool, but int is expected

Constructing trees with positions:

* obtain position from lexer when parsing beginning of tree node
* save this position in the constructed tree
* can also save end positions

What is important is to save information for *leaves*

* information for other nodes can be approximated using information in leaves

**Question:** how would we build a support for IDE that displays, for each identifier, the content of the line where it was declared?