

Homework # 2

due 2008/2/6

Reminder: You may work on homework with other people, but (1) indicate with whom you worked and (2) please use your own words when writing up the results.

1 Measuring a Scanner

On a “grid” machine (or any machine using gcc 3.1 or up) run the C++ preprocessor (`g++ -E`) on the following program:

```
#include <iostream>

int main()
{
    std::cout << "Hello, world!" << std::endl;
}
```

- (a) List the files that the preprocessor must look at in its *input*. (Hint use `g++ -E hello.cc | grep '\# 1'` to find the lines where an included file starts in the result. Ignore the first three lines of output.) Use `wc` to estimate how many characters the preprocessor must look at, broken down by file. (Some files are read twice; don't forget to add up *all* the characters read.)
- (b) How many characters are in the *output* of the preprocessor?
- (c) The output of the preprocessor is the input of the scanner proper. Regular comments have already been removed, but the preprocessor adds lines starting with '#' that are treated as comments by the scanner. How many tokens (approximately) represent the *output* of the scanner? (Please indicate how you made the estimate.)
- (d) Write a paragraph analyzing your results. What do they indicate about scanning (even short) programs in C++?

2 Regular Expressions

Solve Exercise 2.3(b) on page 99 of the textbook (2.1(b) on page 96 of new textbook). Make sure your solution accepts **(**)** and **(***)** (but not **(*)** or **(* *) (*)**).

Turn in both the automaton and the regular expression along with your explanation of how they handle the test cases given above.

3 Lexical Errors

As a programmer, what would you like a scanner do when it encounters errors such as

- a character that cannot legally occur in any token except a string or a comment: (such as `$` in `Cool`),
- a badly formed token, such as `_x` in `Cool`,
- a badly terminated string or character?

Give sound reasons for your opinions. What does the Cool compiler do in such cases? Does it correspond with what you would like? (Please give actual examples of running `coolc`.)