

## Homework 3 (due 2/6)

1. (20 pts.) Write a complete C++ program to output the top 20 most frequently used words in a text file, in sorted order, starting with the highest frequency. The file name should be read from input. If the file has only  $k < 20$  different words, then only  $k$  entries are printed.

Example: the output for `short.txt` (in homework directory) is

```
2 this
2 test
2 is
1 another
1 a
```

Example: the output for `anna_karenina.txt` (in homework directory) is

```
16266 the
11494 and
9899 to
8442 of
6146 he
5935 a
5482 in
5158 was
5035 his
4661 that
3887 her
3810 had
3643 with
3353 she
3218 not
2827 I
2716 at
2469 for
2363 it
2287 as
```

Example: the output for `don_quixote.txt` (in homework directory) is

```
20758 the
16571 and
13417 to
12744 of
7164 that
6956 a
6801 in
5773 I
5693 he
4537 for
4517 his
4258 as
4193 it
3485 is
3476 with
```

3423 was  
3403 not  
2996 be  
2713 my  
2539 Don

Notes:

1. Use `map<string, int>` to keep track of the frequency count for each word.
2. Each element of `map<string, int>` is a `pair<string, int>`, which has two fields: `first` and `second`.
3. `pair` is declared in the header file `<utility>`.
4. `map` cannot be sorted. To sort the pairs using the `second` field (frequency count), first copy the contents of your `map<string, int>` to a `vector<pair<int, string>>` (note that the order of the fields is reversed) and then sort the vector.

You must turn in by noon of the due date:

- a hard copy of your code with your name, section number, and sample output; and
- send **one** email message with the subject line: HW3 Your\_last\_name Your\_section to `cs60@math.scu.edu` with your code attached. Please do **not** zip your file.