**Project 2: due December 21 2017, 11:30am in Dropbox.**

When talking with students, I noticed many have seen different movies with different opinions of various movies. In this project, we are to take survey on the preference of CS students to movies. The data collected from the survey can be stored in heaps and PriorityQueue for new students to query about the most popular, the most-viewed, highest rating, and the latest movie.

Each student is to collect a movie rating information from a survey to students in our school of computer science. The survey results are stored in a common text file using the format of [www.imdb.com](http://www.imdb.com) rating system. This step gives students opportunities to interact and discuss the movie topic that is popular to general public. Students will practice using a heap to facilitate the selection of highest rating and/or most popular and/or latest movies. First students need to know how to create a heap and use it to implement a priority queue, and then use priority queue with Comparators to select the highest one with various Comparators. The students get to learn about the difference between selecting the maximum and sorting the entire list, as well using different criteria for selection.

1. Collect movie ratings from the students of SCS and add the rating info to the common rating file in the shared Dropbox directory. The format is as follows. Each line is an entry of one movie. Each movie has 5 fields: rating percentage, number of students saw the movie, average rating, movie name, and movie release year. An example is given later. [1 point]
2. Create a movie class with the 5 field information of rating. [1 point]
3. Create comparators to compare 4 fields except movie names. [1 point]
4. Create a heap using various comparators. [1 point]
5. Create PriorityQueue that use the heap to answer users’ query of highest percentage, most-viewed, highest rating, and the latest movie. [1 point]
6. Demo the interactive program to the class.

Appendix:

[www.imdb.com](http://www.imdb.com) rating data format

0000012211 174 8.4 “Star Wars: The Last Jedi” 2017

0000122110 369 7.9 “Thor: Ragnarok” 2017

The first number has 10 digits indicating the percentage of viewers who gave the corresponding numeric score. “0000012211” means that between 0-10% viewers give the scores of 1,2,3,4, and 5, between 10-20% give 6, 9, and 10, and between 20-30% give 7 and 8. Can you explain what does “0000122110” mean? And is it better or worse than “0000012211”?

The second number is the number of students who have seen the movie.

The third number is the average rating of the movie by the students.

The fourth is the movie name and the fifth is the movie release year.