DEPARTMENT OF PHYSICS & ASTRONOMY
3459 EXAM-2
14:00 - 17:00 : 13th December 2010

Please read the exam guidelines, rules, instructions and marking criteria at
`http://moodle.ucl.ac.uk/mod/wiki/view.php?id=13963&page=Final+exam`
(linked from the *Exams and Coursework* page).

This exam is worth 50% of your final mark for the course. The duration of the exam is 3 hours. Students should upload the Java source code files for their solution using Moodle under the section headed "Exam 2".

The GHCND project publishes a collection of daily measurements from weather stations around the world, some dating back to the nineteenth century. In this exam you will be processing some of these data, which are provided in a simplified format.

Three files are provided in the following web directory:
http://www.hep.ucl.ac.uk/undergrad/3459/exam-data/2010-11/.

- The file stations.txt contains a list of weather stations: the first field is the identification code (ID), and the rest of the line is its name. The first two characters of the ID are the identification code of the country where the station is situated.
- The file countries.txt is a list of countries. The first field on each line is the identification code of the country, and the rest of the line is its name.
- The file readings.txt contains measurements of temperatures and precipitation made at some of the stations. Each line starts with the following four fields:
  - the ID of the station where the measurement was made;
  - the year;
  - the month, as an integer number in the range 1–12;
  - the type of measurement:
    - TMIN: minimum temperature
    - TMAX: maximum temperature
    - PRCP: amount of precipitation

  The following 31 fields are integers representing the measurement made on each day of the month. The temperatures are given in units of 0.1 °C and precipitations are given in units of 0.1 mm. Missing measurements, and measurements that would correspond to non-existent dates (e.g. 31st February), are represented by the number -9999.

Part 1: 15/50 marks

Write a program to do the following:

- Read the data from readings.txt and store the TMAX measurements in an appropriate container.
- Read the data from stations.txt and countries.txt and store them in appropriate containers.
- Find the highest temperature measurement that occurs in the data, and for each time it occurs (note, the same temperature may be recorded multiple times and by different stations) print out the date, station name and country name.

Part 2: 20/50 marks Add additional analysis methods to your code to:

- Determine the country (or countries) with the lowest mean maximum temperature.
- For each year in the data sample, determine the country (or countries) with the lowest mean maximum temperature that year.

Part 3: 15/50 marks

Make the following enhancements to your code:

- Define an interface to calculate any statistic (e.g. mean, median, maximum, RMS, etc.) from the maximum temperature data for a particular country and time period, specified in terms of months and years.
- Modify your code so that it uses this interface to determine the country (or countries) with the lowest mean maximum temperature.
- Provide an alternative implementation of this interface which calculates the RMS of the specified data, and determine the country (or countries) with largest variation (i.e. the largest RMS) of the maximum daily temperature in the year 1966.

Uploading your work

If you use your own classes from earlier modules, make sure you upload them as well as any new classes you create during the exam.

END OF PAPER