CS23710 Assessed Assignment 2014-2015 "Counting the Cetacean Mammals of Cardigan Bay".

David Price and Fred Long

November 8, 2014

1 Introduction

PLEASE NOTE: We require that your programs are compliant with a recognised version of ANSI 'C'.

You **must** use NetBeans and the gcc compilers and gdb debuggers accessible from it to develop and debug your program. In addition, we also specify that you *MUST* use **remote build host support** and you must select to use `central.aber.ac.uk` as the remote build host.

Our "official environment" in which all your user programs must run, and in which we will test them, is that provided by the environment described above.

2 Monitoring Dolphins and Porpoise between Strumble Head to Braich-y-Pwll

You have been contracted to produce a software system to assist marine biologists who are interested in monitoring the cetacean mammals of Cardigan Bay. While several species of cetacean mammals are present in the bay, this project focusses on Harbour Porpoise and Bottlenose Dolphins.

The sea area is depicted on the Eastern section of the Admiralty Chart number 1410 titled "Carnsore Point to Braich-y-Pwll".

3 Background and Simplifications

NOTE: While much of the information supplied in this assignment specification is closely related to real issues of monitoring marine mammal populations, various simplifications have been made to reduce the complexity of the problem.

The observers have been equipped with modern accurate equipment. The equipment allows an observer to give a bearing and a distance to the sighted cetacean mammal and state whether a Harbour Porpoise or a Bottlenose Dolphin has been sighted. The equipment (and the observers) are not perfect and so some errors will exist in the information they report.

Some observers are on shore but others are located on boats at sea.

4 The Programs you must Create

You are required to design, write and test an ANSI C program, with three clearly identifiable separate features, to help solve various tasks that could assist a marine biologist to assess mammal populations.

A firm statement: Your program **MUST** be written so that it is very clear to us which sections of code provide the three features described below.

When running your program, it **MUST** be possible for us to ask it to just run the first feature, then run the second feature and finally run the third feature.

The first feature, specified in more detail in section 8.1, requires that you analyse a set of **sightings** of cetacean mammals and calculate the implied locations (latitude/longitude) of each mammal.

The second feature, specified in more detail in section 8.2, requires that you analyse the calculated locations of cetacean mammals and decide how many mammals have really been sighted. That is, you decide whether some sightings actually refer to the same mammal and if so, calculate **corrected** locations.

The third feature, specified in more detail in section 8.3, requires that you analyse the corrected locations of the cetacean mammals and decide how many separate pods of mammals are present in the bay.

5 Locations and Distances

5.1 Units used in our Data and for Measurements and Calculations

In accordance with one common maritime choice, the "units" used in our data are degrees of latitude and longitude for locations, degrees from true north (clockwise) for directions, and nautical miles for distances.

Observer and Cetacean Mammal Positions
Positions of observers and cetacean mammals will be provided or calculated as a latitude and a longitude in degrees and expressed as real numbers.

Latitude: will be given (or should be calculated) as positive values in the Northern hemisphere and negative values if a location is south of the equator. Thus, latitude goes from 0.000 at the equator to 90.000 at the North Pole and to -90.000 at the South Pole.

Longitude: will be given (or should be calculated) as positive values if the location is to the East of the Greenwich meridian and negative values if the location is to the West of the Greenwich meridian. Thus, longitude will go from 0.000 at Greenwich to negative values as you travel due West and -180.000 is the far side of the world. As you head due East from Greenwich, values increase positively and 180.000 is also the far side of the world, and indeed exactly the same place as -180.000.

Example Locations: Here are two example locations.

1. A latitude of 52.500 with a longitude of -5.500 is in the middle of St. Georges's Channel (i.e., between Wales and Ireland)
2. A latitude of 51.767 with a longitude of 1.283 is the approximate location of the Gunfleet light about 10 miles off Frinton-on-Sea on the UK East coast just North of the entrance to the Thames estuary in the southern North Sea.

Bearings (Directions)
The bearing of a mammal sighting will be given in "degrees true". The direction reports which direction the observer is looking when they sight the cetacean mammal. Thus, a bearing of 45.000 means the observer is looking towards the North East, a bearing of 180.000 means due South and 270.000 means due West.

Ranges (Distances)
The range of a sighting (distance from observer) will be given in nautical miles and expressed as real numbers. [Note: A nautical mile corresponds to the distance equivalent to moving north or south by one minute of latitude, i.e. by one $\frac{1}{60}$ of a degree.]

Our Sea Area
We are only "interested" in what cetacean mammals are doing if they are within "our" sea area. In other words if they are in the sea area covered by the Eastern part of the chart "Carnsore Point to Braich-y-Pwll".

- If sighting reports indicate that a mammal is further South than a latitude of 52.00, or further North than 52.833, you should ignore that report.
- If sighting reports indicate that a mammal is further East than a longitude of -4.000, or further West than -5.500, you should ignore that report.

5.2 Calculating a Cetacean Mammal's Position

You need to calculate the reported location of cetacean mammals based on the sightings that are received.

You are required to write a function that calculates the location of a cetacean mammal. The function should take an observer location, a bearing and a range as parameters and it should produce a calculated location as its return value.

Here's the "mathematics" you need to be able to calculate that position.

In what follows:

OLAT represents an observer's own latitude in degrees

OLATR represents an observer's own latitude converted into radians

OLONG represents an observer's own longitude in degrees

BG represents the bearing from the observer to the cetacean mammal measured from True North in degrees

BGR represents the bearing of the cetacean mammal from True North converted into radians

RANGE represents the range from the observer to the cetacean mammal measured in nautical miles

CMLAT represents a cetacean mammal's calculated latitude in degrees

CMLONG represents a cetacean mammal's calculated longitude in degrees

To convert an angle in degrees to an angle in radians you have to multiply by PI and divide by 180.0 The value for PI can be accessed as a constant called M_PI which is defined in the math.h standard header file.

That is,

(angle in radians) = (angle in degrees) * M_PI / 180.0;

The references to sin and cos in the equations given below are references to the mathematical functions sin and cos.

Pre-compiled versions of these exist in the C standard maths library and their signatures are again defined in the math.h standard header file. As well as including the math.h header file in your programs, you will also need to tell the compiler to link in the math library. This is normally done by having an extra -lm parameter on the compiler command line.

CMLAT = OLAT + (RANGE * cos(BGR)) / 60.0;

CMLONG = OLONG + (RANGE * sin(BGR) / cos(OLATR)) / 60.0;

The accuracy of above mathematics depends on a few simplifying assumptions such as the earth being round and the locations not being too near the North or South poles.

5.3   Calculating Distances Between Locations

For features 2 and 3 you need to calculate the distances between the locations of cetacean mammals.

We have provided a compiled version of one function that calculates the distance between any two locations, that is detailed in a section 7 of this worksheet.

6   The Data Files Provided

We provide a set of data files to describe the location of the observers and the sightings they make of cetacean mammals.

6.1   Observer Locations

The Observer Location files contain a set of lines of information representing the location of our observers at a certain point in time.

Each data file contains a fixed format introduction followed by a set of entries, one per observer. Your program **MUST NOT** assume any fixed specific number of observers but should process the file until end-of-file is encountered.

The first line of the file contains the date and the time using a 24 hour clock. The line contains six integers, separated by spaces, which are day month year hour minute second.

The remainder of each file contains lines of information, each line representing a single observer and their location. You may assume that all the observer ID values are all unique.

The values supplied on each line are separated by spaces and the values and their types are in the following order:

an "ID" for the observer as an alphanumeric string without any embedded gaps;

the latitude of the observer in degrees as described above;

the longitude of the observer in degrees as described above;

Here is the contents of a (small) example file for the 5th November 2014 at 14:53:00.

05 11 2014 14 53 00
AB01 52.408 -4.217
XY23 51.750 -4.300
PQ29 52.100 -6.000
NY23 52.000 -5.900

6.2   Cetacean Sightings

The Cetacean Sightings files contain a set of lines of information representing the cetaceans that have been sighted by the observers.

Each data file contains a set of entries, one per sighting. Your program **MUST NOT** assume any fixed specific number of sightings but should process the file until end-of-file is encountered.

The file contains lines of information, each line representing a single sighting, the observer's ID, the cetacean type, the bearing to the cetacean mammal and its range.

Note: an observer might sight multiple cetacean mammals and multiple observers will often sight the same cetacean mammal.

The values supplied on each line are separated by spaces and the values and their types are in the following order:

an "ID" for the observer as an alphanumeric string without any embedded gaps;

the type of cetacean mammal that has been seen as a single letter (P for Harbour Porpoise, D for Bottlenose Dolphin);

the bearing from the observer to the cetacean mammal measured from True North in degrees as described above;

the range from the observer to the cetacean mammal measured in nautical miles as described above;

Here is the contents of a (small) example sightings file.

AB01 P 270.0 4.10

AB01 D 275.0 4.50

XY23 D 320.0 6.79

XY23 D 340.0 7.50

PQ29 P 230.0 6.50

NY23 D 359.0 8.30

You will notice that some observers have sighted more than one cetacean mammal.

6.3 Alternative Data Files

The data we provide must not be "hard coded" into your program in any way.

It should be possible for us to provide new data files and then use your program to process the data purely by us providing the names of the new files.

Indeed, we may choose to deliberately have some extra data files, of exactly the same format, which we choose to use when we evaluate your program, which quite deliberately are **NOT** available to you during development!

While any files we would use would be of the same format to those described above, we would have different numbers of observers, cetacean mammals etc.

7 Pre-compiled Program Code that we Supply

We are providing you with a pre-compiled version of a function that you are **required** to use.

1. The function calculates the shortest distance, in nautical miles, between two places on the earth's surface. It thus calculates the so-called Great Circle distance. The code is based on the calculation as described on the web page

    http://en.wikipedia.org/wiki/Great-circle_distance

    The detailed definition of the interface to our function is provided as a well commented C header file which is downloadable from the Blackboard course web site.

    However, in simple terms, it expects you to supply two locations and it then calculates the distance between them.

    Each "location" must be supplied as a struct with two fields. The fields are real numbers (as type double) and represent the latitude and longitude of a location. The function's return value is a double and is the distance between the two locations in nautical miles.

8 Your Missions

8.1 Feature 1 : Your Initial Mission

You are required to have the following features in your program.

1. The program must prompt for the name of a file which contains details of all the observer locations that are active in our sea area.

   Your program must open the file, read in all the data and store it in a suitable set of data structures and variables etc.

2. The program must prompt for the name of a file which contains the sightings of cetacean mammals.

   Your program must open the file, read in all the data and store it in a suitable set of data structures and variables etc.

3. Your program should now calculate the locations of all the cetacean mammals as sighted by the observers.

4. Your program should display the locations of the cetacean mammals, their types and which observer sighted them in a suitable tabular format.

8.2 Feature 2 : Your Main Mission

For this feature you must analyse the locations of the cetacean mammals that have been sighted and decide which reports refer to the same animal.

Two (or more) sighting reports are considered to refer to the same mammal if the calculated locations are within 0.02 of a nautical mile (approximately 40 metres) of each other.

You should set the "true" location of the cetacean mammal to the average of the multiple sightings.

You must present your results in a suitable tabular form which, for each cetacean mammal, should include the individual multiple sighting locations as well as the averaged location and the type of the mammal.

8.3 Feature 3 : Your Pod Mission

Cetacean mammals often swim in groups referred to as pods. For the purposes of our program, a cetacean mammal is considered to be in a pod if it is within 0.1 of a nautical mile of any other cetacean mammal within a pod.

For this feature you must calculate how many pods of cetacean mammals are active in our area at the time of the sightings. For each pod you must display how many animals are within the pod and the location of each of the mammals. You must present your results in a suitable tabular form.

8.4 Implementation Requirements

You are expected to make good use of the facilities of the C programming language (1989), C99 or C11 that are supported by our compilers.

Your program **must** make use of **arrays** and/or linked data structures at appropriate places.

Your program **must** make good use of **functions** to modularise your code in a sensible way.

8.5 Comments and layout

We do of course expect you to sensibly comment your programs, making sure that all comments add real value and do not just, in essence, duplicate code. The program should have good layout and must use meaningful names for variables, structures and other identifiers.

8.6 Choices you are free to make

You need to read in the data files and store the information within suitable data structures, which you have designed, within your program.

Your program should use multiple functions, located in multiple files as you choose.

As part of our assessment, we will evaluate the quality of the choices you have made.

8.7 Testing and Data Processing Requirements

As well as any testing you conduct to debug and validate your program, you **MUST** also analyse each of our data files.

We will produce a second **"further information" document** clearly stating exactly how many datafiles we have produced, their names and how you access them.

9 Assessment Criteria

Bearing in mind that this project concerns software development, the most appropriate "assessment criteria" are those in Appendix AA of the student handbook, namely those for "Assessment Criteria for Development".

This assignment is worth 50% of the total marks available for this module.

NOTE: Solutions which do not tackle the pod mission could potentially be awarded a mark of up to 75% if they cannot be criticized in any manner and are accompanied by all other materials as indicated below.

Solutions must also tackle the pod mission for marks of 76% or above.

10 The Material You Must Submit

The submission must be made electronically using Blackboard. There will be appropriate links in the Assignments part of the course materials for module CS23710.

You must submit (via Blackboard) two things.

1. NetBeans Project Export.

   Your NetBeans Project Export **MUST** be made using NetBeans. You should make it very clear which version of NetBeans has been used and on what computing platform. A stated earlier, your **MUST** use remote build host support and use `central.aber.ac.uk` as the build host.

   You MUST make sure that your export includes all of the source code of your program addressing each of the three required features. **Note:** we will be judging the quality of the layout of the program you have produced.

2. A Design, Compile / Run Document

   You MUST submit a document (in PDF format). This document must include:

- A short section of between two and four sides of A4 in length describing the design decisions you made. Include diagrams if you wish to help explain your design.
- If you have used any minor pieces of code taken from elsewhere, or drawn understanding from published books or web pages, you should include appropriate references and a bibliography in your document. Please make sure you fully understand the department's stance on plagiarism as described in the current student handbook.
- The document must also include a section showing the output produced when attempting to compile your program and any errors or warnings that are produced. This can be a collection of "screen shots'
- The document must also include a section showing the output produced when running your program for each of the three features and it processing our supplied data files. This can be a collection of "screen shots".

NOTE: This is an "individual" assignment and must be completed as a one person effort by the student submitting the work.

By making your submission on Blackboard it will be deemed that you are declaring the submission to be your own work.

11   Submission Date

Your solution to this assignment must be made on AberLearn Blackboard before 3:00pm on Friday 12th December 2014.

It is not acceptable to miss any lecture for the purposes of completing assignment work, in terms of the department's monitoring of student attendance and any actions that might be taken in respect of poor attendance.

12   Feedback

To enable you to generate "self feedback" we plan to release our own "outline/sample solution" (which we note will not be perfect!) during December, hopefully before Xmas day. You will then be able to use this to see how we tackled the problem, especially in any areas where you had difficulties.

It is our intention to provide you with individual feedback on your work by approximately mid January.