FB 08 Informatik
Lehrstuhl f. Software-Technik und Bioinformatik

Prof. Dr. Andreas Hildebrandt
Dr. Marco Carnini

BIOINFORMATICS II - SS 16

1. EXERCISE SHEET

TO BE DELIVERED NOT LATER THAN 01-05-2016

	Exercise	Points
Theoretical	1	10
Practical	2	10
Practical	3	10

Exercise 1: Understanding PDB file (10 Points)
The PDB file format is the most commonly used format for storing information about the 3D structure of proteins. The specifications of the format can be found at
`http://tinyurl.com/jeg3eed`
Reading the documentation, answer the following:

(a) What will be the record to search for water molecules in the PDB?

(b) Which columns should be read if we want to extract the element symbol of an atom?

(c) Which columns contain the x, y, z coordinates of an atom?

(d) Which columns contain the residue name (amino acid)?

(e) Which column contain the chain name?

Exercise 2: Parse a PDB file – C++ (10 Points)
Using simply `C++` (not using `Ball`), create a program that accept as input parameter the name of a file, then open it, and write only the coordinates of atoms. Consider for testing your program the file `1l2y_simplified.pdb`; it is not a real PDB file, as all the PDB file names are 4 characters long (not including the extension `.pdb`) and it was created for this exercise. If the program work properly, the output should start with the following lines (more spaces allowed):

```
-8.901   4.127  -0.555
-8.608   3.135  -1.618
-7.117   2.964  -1.897
-6.634   1.849  -1.758
-9.437   3.396  -2.889
-10.915   3.130  -2.611
```

The last line should be:

```
2.831  10.040   2.676
```

Exercise 3: Handling a PDB file – C++ (10 Points)
Create a `C++` (not using `Ball`) code that, given as input a filename, read the pdb and translate all the atoms (i.e., add to all the coordinates) the vector **d**:

$$\mathbf{d} = (1, 1, 1)$$

For example, given the vector **x**

$$\mathbf{x} = (1, 2, 3)$$

the translation will be:

$$\mathbf{x} + \mathbf{d} = (2, 3, 4).$$

The result expected is the PDB with the other lines unchanged, and only the positions modified in order to get a valid PDB.