NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

School of Science

Information Technologies in Medicine and Biology

Direction: *Bioinformatics*

Biomedical Databases

Postgraduate Student: *Begetis Nikolaos*

Professor: *Gunopoulos Dimitrios*

Deadline Date: *end of May 2013*

Final Project

As we discussed in class I am assigned to create a **text reader** which should find and highlight in the text given each time the names of proteins and when clicking a name the PDB 3D model should appear in screen (**3D protein viewer**) provided with lots of features so that this kind of implementation should be useful for users.

More specifically, we are going to create an Android application which should have two main components. The first one may be a text reader, for instance a PDF viewer, which when given a PDF file it should search for keywords (protein name, PDB name, etc.) and highlight them using them as hyperlinks to the second main component input. The highlighting will be done by some indexing (for efficiency) in a group of words (at first time), and in future work (hopefully, provided in this final project) to all the existing protein names, PDB names, etc, by fetching them from the internet.

As a second component there will be implemented an 3D protein viewer, which each time will be initiated from a "tapping" in the first component's keyword. This "tap" should hyperlink the keyword tapped with a value name (using hash function) and give this value as an insertion to PDB (etc.) webservices so that the 3 dimensional structure will be fetched and represented in our mobile or tablet. The features functionalities that will be provided with the 3D modeling will be:

- Rotatation/Translation/Zooming of the model by finger
- Representations using:
    - Lines
    - Sticks
    - Spheres (van der Waals radius)
    - Alpha carbon traces
    - Thick ribbons

    - Thin ribbons
    - Strands
    - B factor tubes
    - Nucleic acid ladders
    - Nucleic acid lines
    - Solvent 'stars'
- Smoothing of beta sheets
- Coloring:
    - By chain
    - By secondary structure (when defined in SHEET/HELIX records)
    - By Elements
    - Gradation (chainbow)
    - Temperature factor (B factor)
    - Polar/Nonpolar
- Crystallographies:
    - Using unit cell display
    - Showing crystal packing (when defined in REMARK section)
    - Displaying biological assembly (when defined in REMARK section)