Lezione 7

Bioinformatica

Mauro Ceccanti[‡] e Alberto Paoluzzi[†]

[†]Dip. Informatica e Automazione – Università “Roma Tre”
[‡]Dip. Medicina Clinica – Università “La Sapienza”

BioPython

- Installing and exploration
- Tutorial
- First Course Project
- First Start
- First Start with Biopython

Contents

BioPython

Biopython is a set of freely available tools for biological computation written in Python by an international team of developers.

- The web site provides an online resource for modules, scripts, and web links for developers of Python-based software for life science

- BioPython makes it as easy as possible to use Python for bioinformatics by creating high-quality, reusable modules and scripts

BioPython

Biopython is a set of freely available tools for biological computation written in Python by an international team of developers.

- The web site provides an online resource for modules, scripts, and web links for developers of Python-based software for life science

- BioPython makes it as easy as possible to use Python for bioinformatics by creating high-quality, reusable modules and scripts

BioPython

Biopython is a set of freely available tools for biological computation written in Python by an international team of developers.

This wiki will help you download and install Biopython, and start using the libraries and tools

Download

Current Release – 1.52

http://biopython.org/wiki/Download

Biopython installation

Short version

- installer for windows: download Python-2.6.2.msi
- standard install on MacOSX, Linux and Unix:
  - download the source
  - from command line in a terminal:

```
> python setup.py build
> python setup.py test
> sudo python setup.py install
```

Biopython installation

Short version

- installer for windows: download Python-2.6.2.msi
- standard install on MacOSX, Linux and Unix:
  - download the source
  - from command line in a terminal:

```
> python setup.py build
> python setup.py test
> sudo python setup.py install
```

Biopython installation

Short version

- installer for windows: download Python-2.6.2.msi
- standard install on MacOSX, Linux and Unix:
  - download the source
  - from command line in a terminal:

```
> python setup.py build
> python setup.py test
> sudo python setup.py install
```

Biopython installation

Short version

- installer for windows: download Python-2.6.2.msi
- standard install on MacOSX, Linux and Unix:
  - download the source
  - from command line in a terminal:

```
> python setup.py build
> python setup.py test
> sudo python setup.py install
```

Biopython installation

Long version

- http://biopython.org/DIST/docs/install/Installation.html

Biopython installation

Best version ;-)

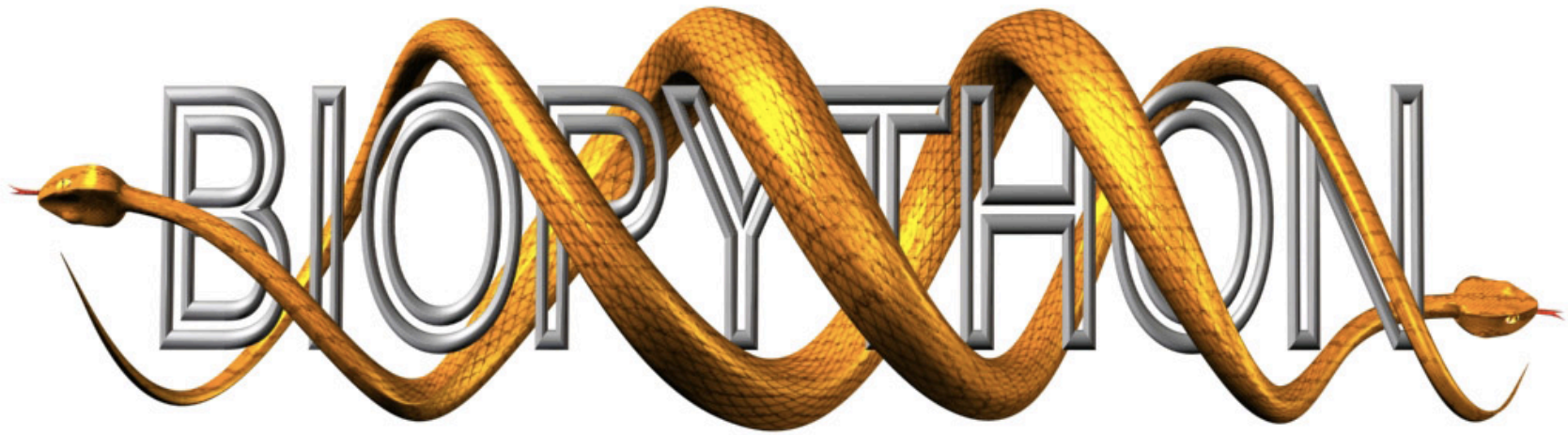
- from a terminal, with easy_install package already installed:

```
> easy_install -f http://biopython.org/DIST/biopython
```

Contents

Biopython Tutorial and Cookbook

by Jeff Chang, Brad Chapman, Iddo Friedberg, Thomas Hamelryck, Michiel de Hoon, Peter Cock, and Tiago Antão



Contents

Biopython Tutorial and Cookbook

Contents

PDB: Atomic Coordinate Entry Format Description

Learn to parse PDB files, locally and on the web

Protein Data Bank Contents Guide

- Introduction
- Title Section
- Primary Structure Section
- Heterogen Section
- Secondary Structure Section
- Connectivity Annotation Section
- Miscellaneous Features Section
- Crystallographic and Coordinate Transformation Section
- Coordinate Section
- Connectivity Section
- Bookkeeping Section

Contents

First project

Curation of records of PDB files of aminoacids

Start from:

Amino Acids web page

and

Library of 3-D Molecular Structures, in particular from Amino Acids Section

Contents

Start using Biopython

Importing the package

```
Python 2.6.3 (r263:75184, Oct  2 2009, 07:56:03)
[GCC 4.0.1 (Apple Inc. build 5493)] on darwin
Type "copyright", "credits" or "license()" for more
    information.

IDLE 2.6.3
>>> import Bio
>>> print Bio.__version__
1.51
>>>
```

Start using Biopython

`help()` on the package

```
>>> help(Bio)
Help on package Bio:

NAME
    Bio - Collection of modules for dealing with
        biological data in Python.

FILE
    /Library/Frameworks/Python.framework/Versions/2.6/
        lib/python2.6/site-packages/biopython-1.51-py2.6-
        macosx-10.3-fat.egg/Bio/__init__.py

DESCRIPTION
    The Biopython Project is an international
        association of developers
    of freely available Python tools for computational
        molecular biology.

    http://biopython.org
```

Start using Biopython

`import` the PDB package

```
>>> from Bio.PDB import *
>>> dir()
['AbstractPropertyMap', 'Atom', 'Bio', 'CaPPBuilder', '
    Chain', 'DSSP', 'Dice', 'Entity', 'ExposureCN', '
    FragmentMapper', 'HSExposure', 'HSExposureCA', '
    HSExposureCB', 'Model', 'NeighborSearch', '
    PDBExceptions', 'PDBIO', 'PDBList', 'PDBParser', '
    PPBuilder', 'Polypeptide', 'Residue', 'ResidueDepth',
     'Select', 'Selection', 'Structure', '
    StructureAlignment', 'StructureBuilder', '
    Superimposer', 'Vector', '__builtins__', '__doc__', '
    __name__', '__package__', 'calc_angle', '
    calc_dihedral', 'extract', 'get_surface', 'is_aa', '
    m2rotaxis', 'make_dssp_dict', 'mmCIF', '
    parse_pdb_header', 'refmat', 'rotaxis', 'rotaxis2m',
    'rotmat', 'standard_aa_names', 'to_one_letter_code',
    'vector_to_axis']
```

Start using Biopython

`help()` on the package

```
>>> help(Bio.PDB)
Help on package Bio.PDB in Bio:

NAME
    Bio.PDB

FILE
    /Library/Frameworks/Python.framework/Versions/2.6/
        lib/python2.6/site-packages/biopython-1.51-py2.6-
        macosx-10.3-fat.egg/Bio/PDB/__init__.py

DESCRIPTION
    Classes that deal with macromolecular crystal
        structures. (eg.
    PDB and mmCIF parsers, a Structure class, a module
        to keep
    a local copy of the PDB up-to-date, selective IO of
        PDB files,
    etc.). Author: Thomas Hamelryck.  Additional code by
```

Start using Biopython

`help()` on the module `Atom`

```
>>> Atom
<module 'Bio.PDB.Atom' from '/Library/Frameworks/Python.
    framework/Versions/2.6/lib/python2.6/site-packages/
    biopython-1.51-py2.6-macosx-10.3-fat.egg/Bio/PDB/Atom
    .pyc'>
>>> from Bio.PDB.Atom import *
>>> help(Bio.PDB.Atom)
Help on module Bio.PDB.Atom in Bio.PDB:

NAME
    Bio.PDB.Atom - Atom class, used in Structure objects
        .

FILE
    /Library/Frameworks/Python.framework/Versions/2.6/
        lib/python2.6/site-packages/biopython-1.51-py2.6-
        macosx-10.3-fat.egg/Bio/PDB/Atom.py

CLASSES
```

Start using Biopython

Look the atom.py file ...

```
# Copyright (C) 2002, Thomas Hamelryck (thamelry@binf.ku
    .dk)
# This code is part of the Biopython distribution and
    governed by its
# license.  Please see the LICENSE file that should have
    been included
# as part of this package.

# Python stuff
import numpy

# My stuff
from Entity import DisorderedEntityWrapper
from Vector import Vector

__doc__="Atom class, used in Structure objects."

class Atom:
    def __init__(self, name, coord, bfactor, occupancy,
```