

MaMEA User's Guide

Lukas Turcani

November 18, 2016

Contents

1	Introduction	2
2	Installing MaMEA	2
3	Input Files	2
3.1	Format and structure.	2
3.2	Valid parameters.	3
4	Running MaMEA	3
4.1	From the command line.	3
5	Output	3
6	MaMEA as a Library	3

1 Introduction

MaMEA (MacroMolecular Evolutionary Algorithm) is a genetic algorithm (GA) for chemistry. It aims to be as general as possible, being suitable for use with any class of molecules. While it does not support all types of molecules out of the box it is designed to be easily extended. This is made easier by being written in Python. For notes on how to extend MaMEA to a new class of molecules or add GA operations, see the developer's guide.

2 Installing MaMEA

3 Input Files

To run MaMEA you need to submit an input file. These contain all the details required to start a GA calculation. Among other things, this may include things like which the fitness function you want to use, the number of generations MaMEA should make and how big the population should be.

Using an example, the format and structure of input files is explained in section 3.1. A guide to finding out what GA operations and tools are available is in section 3.2.

3.1 Format and structure.

An example input file can be downloaded from https://sabercathost.com/1aoC/mamea_input.in.

Any lines consisting of only whitespace or beginning with “#” are ignored. This is used to create comments in the input file and allows us to divide it into sections. Notice that there are two kinds of non-comment line in the input file. The first kind has the form

```
keyword; function_name; param1=val1; param2=val2; param3=val3
```

Lines 5, 11, 17 in the example input file are examples of this. Here is line 11 (in case you did not download the file).

```
generational_select_func; stochastic_sampling; use_rank=True
```

This means that in line 11 the keyword is “generational_select_func”, the function name is “stochastic_sampling”, parameter 1's name is “use_rank” and its “True”. As you can see, in this line there is only 1 parameter name-value pair.

Lines with this format define functions used by MaMEA. These functions will correspond to the various GA operations, such as crossover, mating and selection or to things like the fitness function or optimization function (more on this later). The words on these lines are separated from the adjacent ones by a semicolon. Notice that the last word on the line is not followed by a semicolon. The keyword defines which GA operation the line describes. For example, in line 11 the keyword “generational_select_func” signifies that the line describes which

function is to be used for selection members of the next generation. In line 17 the keyword “parent_select_func” signifies that the line defines the function to be used for selecting parents for crossover. A list of valid keywords is described in section 3.2.

The parameter “function_name” is the name of the function which is to carry out the role designated by the keyword. This must correspond to the name of a function defined within MaMEA.

3.2 Valid parameters.

4 Running MaMEA

4.1 From the command line.

5 Output

6 MaMEA as a Library