

Problem Solving Project: Evolutionary Reconstruction of a shredded picture

Pascal Bouvry, Emmanuel Kieffer

1 Objectives

- Introduce a meta-heuristic framework to students.
- Familiarize students with the concept of Genetic Algorithm through a hand-on experience.

2 Introduction

In this project, students will be tasked with restoring the provided shredded picture. The restoration is to be done through the use of Genetic Algorithm (GA). The essential functions for image processing and GA structure are provided in Python. Students are encouraged to study the document of “Distributed Evolutionary Algorithms (DEAP)”, a meta-heuristic framework ¹ since some parts of the code need to be implemented and/or modified, to achieve the result.

In order to restore a shredded documents, one need to put document strands in a correct order as can be seen in the Fig.1. In this project, students are provided with a digital image which is supposedly shredded by shuffling the order of column in the picture. As an end result, the shredded pictures should be restored.

3 Project Description

In this problem, students will be asked to get familiar with a python evolutionary computing library which provide all necessary tools to implement a genetic algorithm. The library offers most of the basic implementations, but students are still required to implements their own functions to make it suit this problem. The DEAP library is very well documented and provide numerous examples. The goal of this project is to make use of the notions learned during this semester as well as taking advantage of the existing tools to solve optimization problems.

The task is to reconstruct a picture that has been shredded. In fact, the columns of the original image have been shuffled so that the objects are not recognizable. To reconstruct the real order, you will have to implement a python genetic algorithm using the DEAP library. In order to evaluate your solution, you will have to your disposal an “oracle” which is a black box function. When you wish to evaluate your solution, you have to provide a permutation list of the columns to the oracle which will give you back a score based on how close the restored picture resemble the original. The closer to 0 the better. The oracle can also be used to display the restored picture. An example is shown in Fig. 2 and Fig. 3 ².

4 Required Tools

The provided source code and DEAP library are implemented in Python. There are also dependencies that need to be installed for the provided source code to work. Since the code is implemented in Python, they are generally cross-platform. In this section, package installation instructions are given.

- Programming language: Python 2 or 3

¹<http://deap.readthedocs.io/en/master/>

²both pictures are taken from <https://github.com/robinhouston/image-unshredding>.

6 SEP 19

S C R E T 100-150000 SEP 72 STAFF
CDE: LONDON 1758Z
TO: DIRECTOR INFO TO IRAN.
WINTEL AJAJA SDPREX INT INTEL

1. ON 4 SEPTEMBER, ADDRESSEE ASKED SDPRETEXT/1 ABOUT WHAT HE HAD LEARNED ABOUT THE ORGANIZATION OF A PRO-BAZHTIAR MOVEMENT. SDPRETEXT/1 SAID THAT THEIR ORGANIZATION HAD BEEN DONE HAD BEEN DONE BY HOMIZADEH AND A FEW OTHERS. THEY HAD BEEN DEVOTING ALL OF THEIR TIME TO ORGANIZING THE YOUNGER IRANIANS IN GREAT BRITAIN. SDPRETEXT/1 SAID THAT THEY ARE FAMPERED BY THE DADI OF FUNDS. HE SAW THAT TO DATE NO BIG GUNS HAD COME FORTH TO DECLARE FOR OR AGAINST SUPPORT BAZHTIAR.

2. SDPRETEXT/1 SAID THAT TO DATE ALL THE ACTIVITIES IN SUPPORT OF BAZHTIAR WERE PROPAGANDISTIC IN NATURE AND NOT PARTICULARLY EFFECTIVE OR WIDESpread. HE BROUGHT THAT IF THE JEWISH WITHIN MERE AVAILABLE, BAZHTIAR WOULD BE ABLE TO ORGANIZE SOMETHING BECAUSE HE HAS A REPUTATION FOR BEING A GOOD ORGANIZER. SDPRETEXT/1 WAS UNABLE TO STATE WHETHER THE LACK OF EFFECTIVE ORGANIZING ACTIVITIES IN ENGLAND WAS BECAUSE OF A LACK OF FUNDS, OR INDIFFERENCE ON THE PART OF THE IRANIAN COMMUNITY.

3. WITH REGARD TO THE MILITARY, SDPRETEXT/1 SAID THAT HE FRIGHT THAT THE IRANIAN MILITARY WAS TOO MUCH UNDER THE THUMB OF MUSILLAS TO ORGANIZE A COUP AGAINST BAZERGAN. HE HAS HEARD THAT MADANI IS GOING TO BE MOVED BACK TO TEHRAN FROM KHUZISTAN, BECAUSE BOTH THE GOVERNOR AND HIS MILITARY LEADER GIVES HIM AN EFFECTIVE ORDER TO MOVE AGAINST THE GOVERNMENT IF HE SO CHOOSES. SDPRETEXT/1 DID NOT THINK MADANI WAS DOING ANYTHING. IT WAS JUST A QUESTION OF THE GOVERNMENT ACTING WITH CAUTION.

4. SDPRETEXT/1 SAID THAT HE SEES A LOT OF FOOLY AND MERELY LISTENS. HE IS DEFINITELY STATING ON THE MIDLINES WITH REGARD TO THE BAZHTIAR BOMBLT. HE SAID THAT, FROM HIS OBSERVATIONS, IF BAZHTIAR ACCEPTS SUPPORT FROM FORMER ASSOCIATES OF THE SHAH, ETC., IT WILL BE A TACTICAL MISTAKE BECAUSE SHAH IS STILL ROUNDY. THIS QUARTER WOULD BE A MISTAKE.

Figure 1: A reconstructed shredded document



Figure 2: A Shredded picture

Figure 3: A restored picture

- Evolutionary Computing library: Distributed Evolutionary Algorithms (DEAP) library

If you do not have already the pip package manager, go on this page <https://pip.pypa.io/en/stable/installing/> and download get-pip.py file. In a terminal, install pip by entering the command python get-pip.py.

Once pip has been installed, execute all the following commands:

- pip install numpy
- pip install Image
- pip install deap

5 How to use the Oracle?

The oracle is a python binary module named blackbox.py. Fig. 4 shows how to communicate with the oracle.

```
Python 3.6.9 (default, Nov  7 2019, 10:44:02)
Type 'copyright', 'credits' or 'license' for more information
IPython 7.9.0 -- An enhanced Interactive Python. Type '?' for help.

In [1]: 
In [1]: 
In [1]: 
In [1]: 
In [1]: import blackbox
In [2]: oracle=blackbox.BlackBox("shredded.png","original.png")
In [3]: permutation=list(range(128))
In [4]: oracle.evaluate_solution(permutation)
Out[4]: 38299392
In [5]: oracle.show_solution(permutation)
In [6]: 
```

Figure 4: An example of how to use the provided Oracle.

You oracle module should always be placed in the same directory as your code. In addition, the file *original_picture.png.enc* should be also placed in the same directory.

The oracle has only two methods and only accept a permutation list of integers. The size of any permutation list should 128.

- *show_solution(permuation_list)* which display a picture after applying the permutation
- *evaluate_solution(permuation_list, save_file_name)* which provide a score that you want to minimize

Your genetic algorithm should theoretically employed only the method *evaluate_solution(permuation_list)*" in order to evaluate solution. Once the algorithm stops and give you the best found solution, you can call the *show_solution(permuation_list, save_file_name)* method to visualize or save the result.

6 Report

Along with the implementation, students are required to turn in a report that contains details of their work. The guidelines are given in this section.

Problem Solving Project: Evolutionary Reconstruction of a shredded picture

firstname name (student email)

1. Which encoding is relevant for this problem?

Provide the name and the encoding, you are going to use for this problem. Describe it by providing an example

2. Which cross-over operator are you going to use?

Provide its name. Describe it by providing an example (e.g. a picture) to demonstrate how it works. Does it conserve the permutation property?

3. Which mutation operator are you going to use?

Provide its name. Describe it by providing an example (e.g. a picture) to demonstrate how it works. Does it conserve the permutation property?

4. Which selection operator are you going to use?

Provide its name. Describe it by providing an example (e.g. a picture) to demonstrate how it works. Does it conserve the permutation property?

5. Implementation

Use the genetic algorithm implementation provided by DEAP using the operators and the parameters described above. You have to provide a work-flow diagram to describe how it works. Use the following parameters:

1. population size (n): 100
2. generations ($ngen$): 500
3. cross-over probability ($cxbp$): 0.9
4. allele mutation probability ($indpb$) : 0.005
5. mutation probability ($mutpb$) : 1
6. Additional parameters are left to your discretion (e.g. tournament selection, ...)

6. Experimental Result

Genetic Algorithm (GA) is based on stochastic mechanism, numbers of run need to be performed to obtain a good statistical confidence. You will be asked to perform 30 runs. The best solution of each of the 30 expected runs has to be recorded and all runs should be described in Tab. 1.

In addition, the report should provide average convergence curve. The curve is plot from the average of each generation best results (i.e minimum fitness) from each runs. A record example can be found in Tab. 2. A curve from the table is shown in Fig. 5

Finally, the result picture (best fitness score) must be included in the report whether it is completely restored or partially restored. You can export it by applying the “show_solution(permuation_list,

Table 1: An experimental result table

No. Run	Fitness Value
1	30
2	50
...	...
30	10
Average Fitness	
Maximum Fitness	
Minimum Fitness	
Standard Deviation	

Table 2: A record of population statistics during 100 generations

Generation	Minimum fitness	Maximum fitness	Average fitness	Standard deviation
1	1000	1100	1050	10
2	950	920	935	5
...	
100	480	420	450	2

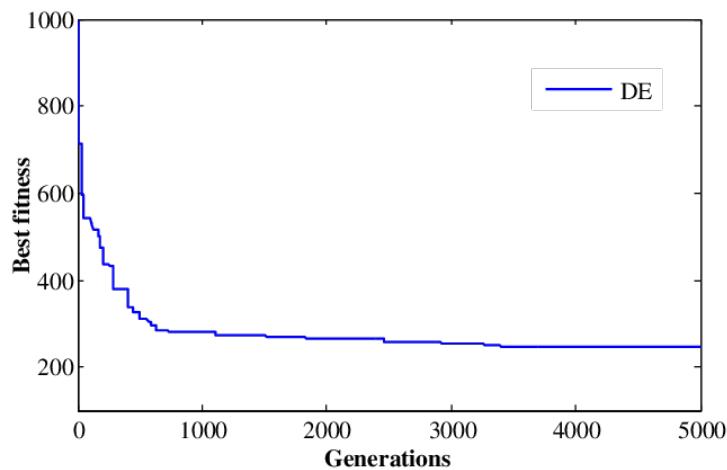


Figure 5: Example of convergence curve with 5000 generations

save_file_name)" method with save the image according to the provided permutation and destination. Ex: oracle.show_solution(permuation,"my_image.png").

Remarks:

Using the knowledge gained during the lectures, you should be able to find all the elements in the DEAP library to perform GA optimization. This is the main goal of this project. Remember that you have to deal with permutation lists with exact size of 64 integers. Negative integers or integers larger than 63 in any list will fail. Last but not least, the problem is very similar as the Traveling Salesman problem seen during the lectures.