Systèmes Complexes et Optimisation Stochastique Massivement Parallèle

Rapport de Projet

Equipe:
ALLEMAND Fabien
LEBOT Samuel

UFR de mathématique et d'informatique
Université de Strasbourg

Contents

List of Figures

1 Introduction

Evolutionary algorithms and genetic programming are used in a wide range of applications. From basic applications in industry and engineering to complex problem resolution in resarch department, these two methods, whose development started with Alan Turing in 1950, no longer need to prove their efficiency.

On the one hand, evolutionary algorithms, inspired of biological evolution, allows us to find quatities or caracteristics of an object, and on the other hand, genetic programming was designed to find interactions between objects. By combining these two methods, it is theoredicaly possible (assuming there is a good way to evaluate our results) to find not only caracteritics defining two entities, but also the way they interact with each other.
These two entities could very well be daily life objects or even animals. We could actually be able to find how birds interact with each other in a flock. These two entites could be less concrete and be ridiculously small: let us find out how electrons interact in matter! But they could also be enormous like celestial bodies...

Let us go back in time and imagine we only have the knowledge and tools to measure the position of the Sun relatively to the Earth. Would we be able to find caracteritics about the Earth or the Sun like their mass or their speed? Is there a way to find out how they interact?

The goal of this project is to utilise the two afore mentionned programming methods to retrieve caracteristics and laws regarding the interactions between two celestial bodies.
We will focus on Earth and Sun as there is a strong interaction between them and yet are comletely different.
First, we will attempt to find caracteristics about the two planets using evolutionary algorithms.
In a second time, we will use genetic programming to re-discover Newton's law of universal gravitation.

2 Sampling

Later in ou study, we will use the trajectory of the Earth in the heliocentric renferential in order to evaluate our models. To be more precise, the score of a model will be the difference between the real Earth's trajectory and the trajectory we estimate with the values or functions we find.
In order to do that, we need to be able to sample the Earth path in that referential. Let us study the Earth in the heliocentric renferential.

System: Earth, assimilated to a material point of mass M_T
Referential: Galilean assumed heliocentric renferential
Coordinates System: Polar coordinates
Balance of forces: Attraction force of the Sun of mass M_S: $\vec{F} = -G\frac{M_T M_S}{r^2}\vec{e_r}$

According to Newton's second law:

$$M_T\vec{(a)} = \vec{F} \iff \begin{cases} M_T(\ddot{r} - r\dot{\theta}^2) = -G\frac{M_T M_S}{r^2} \\ M_T(2\dot{r}\dot{\theta} + r\ddot{\theta}) = 0 \end{cases} \quad (1) \quad \iff \begin{cases} \ddot{r} = r\dot{\theta}^2 - G\frac{M_T M_S}{r^2} \\ \ddot{\theta} = -\frac{2\dot{r}\dot{\theta}}{r} \end{cases} \tag{2}$$

In order to sample the Earth's trajectory, we perform two Euler methods simultaneously: the first one allows us to obtain the speed and the second one gives us the actual position.

$$\begin{cases} r_{n+1} = r_n + \dot{r_n}\delta t \\ \theta_{n+1} = \theta_n + \dot{\theta_n}\delta t \end{cases} \tag{3}$$

$$\begin{cases} \dot{r_{n+1}} = \dot{r_n} + \ddot{r_n}\delta t \\ \dot{\theta_{n+1}} = \dot{\theta_n} + \ddot{\theta_n}\delta t \end{cases} \tag{4}$$

with $\delta t = \frac{365,25\times 24\times 3600}{1024}$ as we want to sample 1024 points.
We also have to set the initial conditions. Let us suppose we start to sample when the Earth is located at the perihelion, as the distance between the Earth and the Sun and the speed of the Earth at this point are known, we obtain the following conditions:

$$\begin{cases} r_0 = 147,1\times 10^9 \\ \theta_0 = \pi \\ \dot{r_0} = 0 \\ \dot{\theta_0} = \frac{2\pi\times 30,2\times 10^3}{\pi\left(3(a+b)-\sqrt{(3a+b)\times(a+3b)}\right)} \end{cases} \tag{5}$$

with $a = 1521,0\times 10^8$ and $b = 1471,0\times 10^8$ the distances between the Earth and the Sun respectively at the aphelion and at the perihelion.
With this set of equations we can write a program that precisely compute 1024 positions of the Earth while orbiting around the Sun.

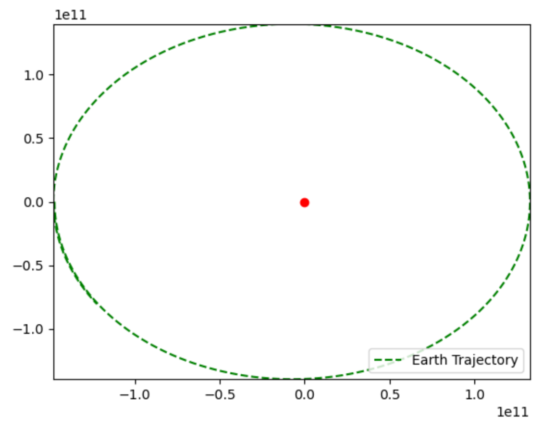

Figure 1: Earth trajectory around the Sun (1024 points sampling)

3 Evolutionary Algorithm

The first part of this project consisted in finding three quantities that influence the orbit of Earth around Sun: the mass of Sun, the mass of Earth and the speed of Earth at the perihelion, using the evolutionary algorithms implemented in EASEA.

In order to do that, we performed three distinct experiences because the sampling method does not allow us to find these three quantities simultaneously. It is impossible to find the mass of Earth and Sun at the same time because they are only bound in a product in Newton's law of universal gravitation. So there are a plethora of pair of values leading to the same result.

The first step consist in sampling the Earth orbit around Sun with the real quantities in order to obtain a list of polar coordinates of Earth in the heliocentric referential.
The genom of each individual correspond to the quantity to find (the mass or the speed) and is initialised to a value of the same order of magnitude as the real value.(ON N'EST PAS CENSE CONNAITRE L'ORDRE DE GRANDEUR...)
From one generation to another, children are created by computing the mean of the two parents. Each individual can mutate with a given probability. This mutation correspond to an increase or a decrease of a given percentage of its genom.
The score is defined as the difference between the real orbit and the orbit with the quantity found in the genoms. That is to say: the sum of the distances between the corresponding points of the two trajectories. By doing so, the result of the algorithm does not directly depend on the value we want to find but rather on samples of coordinates that could have been made without knowing the quantity.
We settled for weak elitism, a form of elitism that ensures the best individual of the global population (children and parents) is conserved in the next generation.

3.1 Mass of Sun

Using the method described above, we managed to find the mass of the Sun with 100% accuracy (Figure 2). As we can see on Figure 3, the trajectory of the Earth around the Sun slowly approach the real world trajectory over generations.

```
================================================================================
SUN MASS: 1988399999999999998650870241689 6.000000 (real),
    19883999999999999986508702416896.000000 (estimated)
SCORE: 0.000000
================================================================================
SUN MASS: 19883999999999999986508702416896.000000 (real),
    20781194869844941721635934699 52.000000 (estimated)
SCORE: 41697618869063.750000
================================================================================
     59          1.430s          6000          6000  0.000000000e+00  3.0e
    +11  3.0e+12  3.0e+13
EASEA LOG [INFO]: Seed: 1675245974
EASEA LOG [INFO]: Best fitness: 0
EASEA LOG [INFO]: Elapsed time: 1.4326
```

Figure 2: Output of the EASEA program

Number of generations	60
Population size	100
Offspring size	100%
Mutation probability	0.2
Mutation variation rate	0.05

Table 1: Parameters used to find the mass of Sun

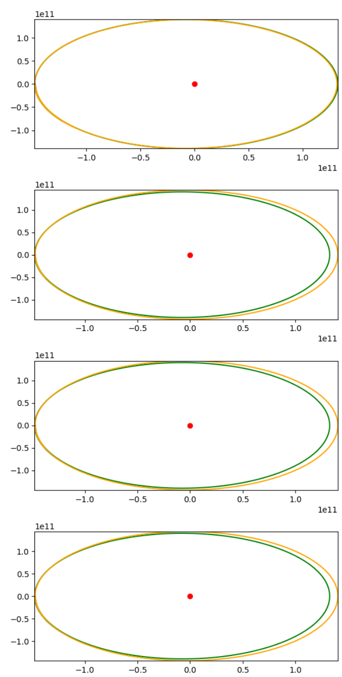

Figure 3: Earth trajectories depending of the mass of Sun

3.2 Mass of Earth

We did not managed to find Earth's mass using evolutionary algorithm as this quantity does not appear in the equations defining Earth's trajectory around Sun. Nonetheless, we have a couple ideas regarding how we could achieve this.
The first being to simply change the referential: use the geocentric referential in order to study the movement of Sun. If we only consider those two planets (Earth and Sun) then their movements are relative to each other, that is to say: Earth is orbiting Sun and Sun is also orbiting Earth. In that case, we can can estimate the mass of Earth exactly the same way as we did in the previous part for the mass of Sun.
Another way to get to that result would be to study those two celestial bodies in a referential centered on the center of Sun's trajectory. But in this case, we would have to study the movement of Earth is a non-galilean referential which is clearly out of the scope of this study.

3.3 Speed of Earth at Perihelion

In order to find the speed of planet Earth at the perihelion we apply the same steps as before: initialisation, crossover, mutation and evaluation. Unless this time we have to change a couples of values. As the speed of Earth in radian per second is so little, it is not well handled by our program (ie: assimilated to zero). To combat that, we simply consider the speed in radian per day leading to values that can properly be handled by computers. This implies also converting the gravitationnal constant and our sampling step to fit the new unit scale.
Once again, we obtained very high accuracy results (Figure 4) and we clearly see how the trajectory evolves throughout the generations (Figure 5).

```
================================================================================
PERIHELION SPEED: 0.017440 (real), 0.017440 (estimated), [1.000000]
SCORE: 0.000000
================================================================================
         99          3.397s           15000           15000 0.000000000e+00 1.6
    e+11 1.3e+12 1.5e+13
EASEA LOG [INFO]: Seed: 1675247174
EASEA LOG [INFO]: Best fitness: 0
EASEA LOG [INFO]: Elapsed time: 3.39933
```

Figure 4: Output of the EASEA pogram

Number of generations	100
Population size	150
Offspring size	100%
Mutation probability	0.2
Mutation variation rate	0.05

Table 2: Parameters used to find the speed of Earth at the perihelion

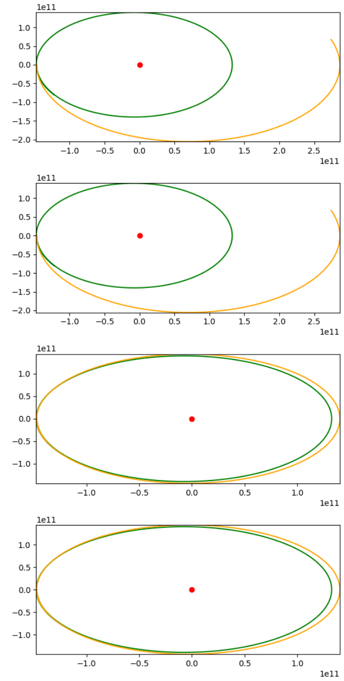

Figure 5: Earth trajectories depending of the speed of Earth at the perihelion