Cemracs 18 Project:

Construction of a Machine Learning tool based on Differential Equations integrated into a dynamic tree allowing biological data assimilation.

Nicolas Bloyet, Hélène Flourent, Emmanuel Frénod

Univ. Bretagne-Sud, UMR6205, LMBA, 56000 Vannes, France

April 13, 2018

1 Introduction

This project enters a research program which goal is to build Machine Learning algorithms based on the discretization of Differential Equations in order to predict biological responses and assimilate biological data. Several such algorithms have already been done.

The Machine Learning algorithm is characterized by 2 facts: it has the capability to predict an output from an input and it can learn how to do it from a Learning Base which is a collection of inputs associated to their outputs. An ODE or a PDE depends on coefficients that can be adjusted to be in accordance with the collections of values that it links. Hence an ODE or a PDE is by nature a Machine Learning tool.

The goal of the project is to explore this idea using the discretization of several ODE or PDE related to one another and integrated into a dynamic structure.

2 Context

Biological reactions prediction is a complex exercise. Indeed biological variations which can be observed, result of multiple and various underlying phenomena - convection, diffusion, fixation, accumulation, delay - difficult to take into account in prediction tools.

In most cases the available data contribute to the increase in the complexity to build biological predictive tool. In fact, biological data contain generally high variability. This variability is due to the individual variability or to the lack for coherence between the experimental trials or to the noise induce by the measure tools. The prediction and simulation tools have to be able to assimilate these complex data.

We can notice that on every scale, living things can be considered as more or less complex dynamic structures. These structures are made of different kind of components which interact and exchange information.

Those kind of data are well represented by *graphs*. A graph is a network of nodes linked by edges, in which both nodes and edges can carry some information.

For example, molecules are structures made of atoms linked each other by covalent bonds. The animals can also be considered as structures made of various organs. These organs receive, treat and redistribute the information which is sent to them, hence they can be considered as *biological functions*.

3 Objectives

The whole living thing, as schematized in figure 1, can be resumed by a network type structure with n nodes (organs). Each of these organs would be composed by :

- a system of PDE, describing the computation core of this organ
- several parameters to adjust, then fixed, used in the PDE
- a variable *state parameter*, used in the diffusion of information

The aim of this study is to focus on the way a message/value is forwarded to other biological functions, maybe assimilating a link strength to a probability, and thus the whole living thing to a bayesian network.

For the least, to state wether or not those models can help to build a challenging multi–dimensional animal representation thanks to graph theory, which could outperform a naive straightforward model.

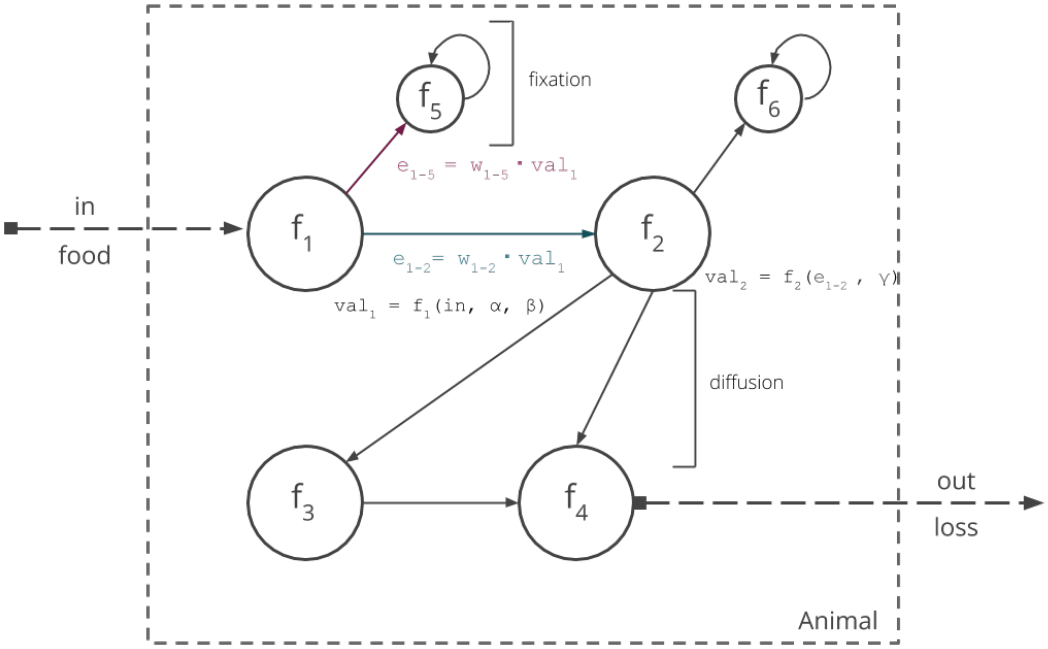

Figure 1: Animal graph representation