Neural cellular automata for modeling protein folding

José Santos[1], Daniel Varela[1]
[1] Department of Computer Science, University of A Coruña, Spain
jose.santos@udc.es

Keywords: protein structure prediction, protein folding, evolutionary computing

Protein folding is the dynamic physical process by which a protein structure assumes its functional shape or conformation and it is a consequence of the interactions through time between the amino acids of the primary structure. It should be treated as a dynamic process and emergent phenomenon, on the contrary to most of the previous research in this computational biology problem, where the intense research was focused on the computational prediction of the final folded structure.

The folding process can be modeled with tools like cellular automata (CA) [1], which can provide and model its emergent and dynamic nature. We are working with CA and machine learning methods to automatically obtain a model of protein folding. CA were implemented with artificial neural networks (neural-CA), instead of classical CA transition rules, incorporating this way the generalization capabilities of connectionist models. Evolved CA decide the most appropriate moves between consecutive amino acids, to obtain the final folded (native) conformation that minimizes its Gibbs free energy (as in the real case).

We have used neural-CA to model protein folding using the basic HP model (Hydrophobic-Polar), with the 2D and 3D lattices [2], more complex lattice models like the Face-Centered Cubic lattice model [3], and the low resolution atomic model of the Rosetta environment [4]. In all cases the importance relies on the automatic modeling of the process, using machine learning and only from known protein structures, on the contrary to the a priori (and non-exact) modeling of protein component interactions followed in molecular dynamic approaches.

References

[1] A. Ilachinski, Cellular automata. A discrete universe, World Scientific, 2001.

[2] J. Santos, P. Villot and M. Diéguez, *Emergent protein folding modeled with evolved neural cellular automata using the 3D HP model*, Journal of Computational Biology **21(11)** 823–845, 2014.

[3] D. Varela and J. Santos, *Protein folding modeling with neural cellular automata using the Face-Centered Cubic model*, LNCS **10337** 125–134, 2017.

[4] D. Varela and J. Santos, *Protein folding modeling with neural cellular automata using Rosetta*, GECCO 2016 ACM Proc. Comp., 1307–1312, 2016.