Memory and dreaming in the Hopfield model.

Enzo Marinari
Sapienza Università Di Roma, Roma, Italy

I will discuss dreaming in Hopfield like neural networks, as a very powerful tool to optimize memory retrieval. The Hebbian unlearning algorithm, i.e., an unsupervised local procedure used to improve the retrieval properties in Hopfield-like neural networks, is numerically compared to a supervised algorithm useful to train a linear symmetric perceptron. I discuss the stability of the stored memories, and notice basins of attraction obtained by the Hebbian unlearning technique are found to be comparable in size to those obtained in the symmetric perceptron, while the two algorithms are found to converge in the same region of Gardner's space of interactions, having followed similar learning paths. A geometric interpretation of Hebbian unlearning is proposed to explain its optimal performances. Since the Hopfield model is also a prototypical model of the class of disordered magnetic system, it might be possible to translate these results to other models of interest for memory storage in materials.

I will also try to investigate more general features of learning and dreaming, and possible approaches to improve dreaming and make it more effective.

References

[1] M. Benedetti, E. Ventura, E. Marinari, G. Ruocco, F. Zamponi, Supervised perceptron learning vs unsupervised Hebbian unlearning: Approaching optimal memory retrieval in Hopfield-like networks, J. Chem. Phys. 156, 104107 (2022).
[2] J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. U. S. A. 79, 2554 (1982).
[3] D.J. Amit, Modeling Brain Functions: The World of Attractor Neural Networks (Cambridge University Press, Cambridge 1989).
[4] P. Peretto, Collective properties of neural networks: A statistical physics approach, Biol. Cybern. 50, 51 (1984).
[5] D.O. Hebb, The Organization of Behavior: A Neuropsychological Theory (Wiley, 1949).
[6] D.J. Amit, H. Gutfreund, H. Sompolinsky, Storing infinite numbers of patterns in a spin-glass model of neural networks, J. Stat. Phys. 55, 1530 (1985).
[7] E. Gardner, Structure of metastable states in the Hopfield model, J. Phys. A: Math. Gen. 19, L1047 (1986).
[8] M. Benedetti, V. Dotsenko, G. Fischetti, E. Marinari, G. Oshanin, Recognition capabilities of a Hopfield model with auxiliary hidden neurons, Phys. Rev. E 103, L060401 (2021).
[9] J. J. Hopfield, D. I. Feinstein, R. G. Palmer, Unlearning has a stabilizing effect in collective memories, Nature 304, 158 (1983).