

Statistical Physics of Computation - Exercises

Emanuele Troiani, Vittorio Erba, Yizhou Xu

September 2024

Week 3

3.1 Lausanne, the happy city

Consider the N citizens of Lausanne as nodes in a given graph G , with vertex set V and edge set E : two people are neighbours on the graph if they know each other. They either are from UNIL or EPFL, which you can consider as a binary variable s_i , $1 \leq i \leq N$ on each of the nodes. The city is a happy one if each couple of neighbours is from different universities.

1. Identify the physical degrees of freedom of the problem, and the couplings. Recall that the degrees of freedoms are the variables in the problem you are free to modify, and the couplings are the constraints on your degrees of freedom. The spins are the degrees of freedom, whose interaction is mediated by the graph (deciding who interacts with whom). Thus the graph defines the couplings of the problem.
2. Write a function of the overall configuration $\{s_i\}_{i=1}^N$ that equals zero if the constraint are not all satisfied. It may be useful to use Kronecker delta functions.

For a pair of adjacent vertices i and j , consider the quantity

$$1 - \delta_{s_i s_j} \quad (1)$$

This quantity will be zero if the nodes have the same value, and one otherwise. We want to impose such a constraint for any edge, so we take the product (as one broken constraint will lead to one factor being zero, hence the whole product to be zero)

$$\prod_{(i,j) \in E} [1 - \delta_{s_i s_j}] \quad (2)$$

where E are the edges in the graph. In physics it's common to use the notation

$$\prod_{\langle i,j \rangle} [1 - \delta_{s_i s_j}] \quad (3)$$

where angular brackets means that we sum over i and j that are neighbours. Here the constraint gives one if the graph is happy and zero if not.

3. Write an equation for the number of configurations of the variables $\{s_i\}_{i=1}^N$ that make the city happy.

One just needs to sum over all the configurations that obey the constraint.

$$\sum_{\{s_i\}} \prod_{\langle i,j \rangle} [1 - \delta_{s_i s_j}] \quad (4)$$

where the product runs over all pairs of neighbours in the graph. Indeed, the constraint will equal one over all happy configurations, and zero otherwise.

4. We define the Garner volume for a fixed graph G as the fraction of configurations $\{s_i\}_{i=1}^N$ that satisfy the constraints. Write an expression for the Gardner volume.

We just need to normalise the previous expression (the total number of configurations that satisfy the constraints) by the overall number of configurations, which is simply 2^N , so the answer is

$$2^{-N} \sum_{\{s_i\}} \prod_{\langle i,j \rangle} [1 - \delta_{s_i s_j}] \quad (5)$$

5. Consider now our graph to be extracted randomly from d regular graph with N nodes. Provide an expression for the average Gardner volume.

The Gardner volume is an exponentially large quantity $\mathcal{O}(e^N)$, so we do not expect it to be self-averaging in the large size limit. This is because there may be graphs G that happen with exponentially small probability, and have a volume which is exponentially far from the one of typical graphs (those that happen with non-vanishing probability). These rare but "far" graphs prevent self-averaging.

The correct self-averaging quantity to look at is the associated entropy density, i.e. $N^{-1} \log$ volume. As it is self-averaging, the entropy density is just the normalised average of the log of the volume, where the average is taken over the graph ensemble (the disorder). We get

$$\frac{1}{N} \mathbb{E} \left[\log \left(2^{-N} \sum_{\{s_i\}} \prod_{\langle i,j \rangle} [1 - \delta_{s_i s_j}] \right) \right] \quad (6)$$

The order in which one averages and takes the logarithm is very important: it's precisely the difference between annealed and quenched averages that we saw in the lecture.

6. Given a graph $G = (V, E)$, compute the total number of constraints and the total number of variables of this constraint satisfaction problem. In which limit do you expect to have a SAT (satisfiable) phase? In which limit do you expect to have an UNSAT (unsatisfiable) phase?

The number of variables is just the number of nodes of the graph, i.e. $|V|$ where absolute value indicates the cardinality of a set. The number of constraints is the number of edges $|E|$, i.e. of neighbouring nodes. We can interpret this problem as a discrete non-linear system of $|E|$ equations in $|V|$ unknowns. Thus, we intuitively expect that for $|E| \ll |V|$ there exist many solutions to our system (underspecified), so we will be in a SAT phase, while for $|E| \gg |V|$ the system will be overspecified and we expect to find no solution (UNSAT phase).

7. Now consider completely different problem: the city is happy if each pair of people that know each other (still defined by a graph G) have a difference of income of more than d . Assume that one can at most earn F francs. What are the degrees of freedom of the model,

and what are the couplings? What is the entropy associated to the Gardner volume for this problem? In which regime will I have a SAT / UNSAT phase?

The degrees of freedom are variables $s_i \in [0, F]$, one for each node. The couplings are again mediated by the graph G . The constraint for each pair to be happy becomes

$$\delta(|s_i - s_j| > d) \quad (7)$$

where now the variables $s_i \in [0, F]$ denote the income of each individual. The volume space is F^N , so the Gardner volume is

$$\mathbb{E} \left[\log \left(F^{-N} \left(\prod_i \int_0^F ds_i \right) \prod_{\langle i,j \rangle} \delta(|s_i - s_j| > d) \right) \right] \quad (8)$$

where the product runs over all pairs of neighbours in the graph.

As above, the number of variables is $|V|$ and the number of constraints is $|E|$, so we expect a SAT phase for $|E| \ll |V|$ and an UNSAT phase for $|E| \gg |V|$.

3.2 Perceptron learning with sign-constrained weights

Consider now the same problem as we saw in the lecture. Given P points ξ^μ in dimension N with labels $\sigma^\mu \in \{\pm 1\}$, we want to find an hyperplane J such that

$$\sigma^\mu \frac{J \cdot \xi^\mu}{\sqrt{N}} > \kappa. \quad (9)$$

We think of the numbers ξ_i^μ , $\mu = 1, \dots, P$, $i = 1, \dots, N$ as being all generated independently from the standard Gaussian $N(0, 1)$. The labels σ^μ are fixed. Also, we take J to have norm $\sqrt{N}$.

Now, suppose that we want to impose an additional constraint, i.e. we require that the coordinates of the hyperplane have fixed signs $\epsilon_i \in \{\pm 1\}$, i.e. that

$$\text{sign}(J_i) = \text{sign}(\epsilon_i). \quad (10)$$

Again, think of the signs ϵ_i as fixed. This is relevant for biologically-inspired problems, where each weight is thought of as a neuronal connection, and some of them are either excitatory ($\epsilon_i = +1$) or inhibitory ($\epsilon_i = -1$).

1. What are the degrees of freedom of the model? And what are the couplings? Note that by degrees of freedom we mean the quantities of the model you can tune to satisfy the constraint, the other quantities that define the constraint are the couplings.

The degrees of freedom are given by the weights J of the perceptron, defining the separating hyperplane. The couplings, i.e. the interactions, are defined by the variables ξ, σ, ϵ , i.e. the data and the sign constraints.

2. Write the averaged entropy density associated to Gardner's volume for this problem.

We have

$$s(\sigma, \epsilon) = \mathbb{E}_\xi \frac{1}{N} \log \int dJ p_N(J) \prod_\mu \theta \left(\sigma^\mu \frac{J \cdot \xi^\mu}{\sqrt{N}} - \kappa \right) \prod_i \theta(\epsilon_i J_i) \quad (11)$$

where $p_N(J)$ is the uniform probability distribution over the N -dimensional sphere of radius $\sqrt{N}$.

3. Argue that the entropy you derived in point 1 does not depend on σ and ϵ .

In the expression above, perform the change of variable $J_i \rightarrow J_i \epsilon_i$ for all i , and then $\xi_i^\mu \rightarrow \xi_i^\mu \sigma^\mu \epsilon_i$ for all μ, i to get

$$s(\sigma, \epsilon) = s = \mathbb{E}_\xi \frac{1}{N} \log \int dJ p_N(J) \prod_\mu \theta \left(\frac{J \cdot \xi^\mu}{\sqrt{N}} - \kappa \right) \prod_i \theta(J_i) \quad (12)$$

where we used that the distribution $p_N(J)$, and the distribution of ξ , are symmetric under any sign flip.