Statistical Physics of Computation 2025 - Exercises

Vittorio Erba, Emanuele Troiani

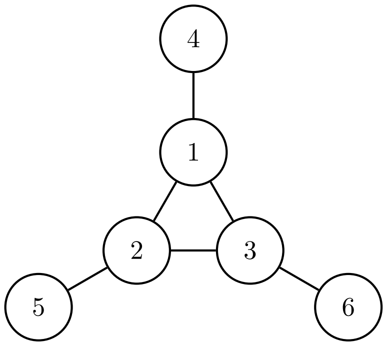
Week 10

10.1 Factor graph representations

1. **Independent set problem**

   The independent set problem is a problem defined and studied in combinatorics and graph theory. Given a (unweighted, undirected) graph $G(V, E)$, an independent set $S \subseteq V$ is defined as a subset of nodes such that if $i \in S$ then for all $j \in \partial i$ we have $j \notin S$. In other words in for all $(ij) \in E$ only i or j can belong to the independent set.

   Example graph:



   (a) Find a bijection between the set of subsets $S \subseteq V$ with $|V| = N$ (where $|V|$ denotes the cardinality of the node set V) and $\{0, 1\}^N$.

   Take $\sigma_1 = 1$ if $i \in S$ and zero otherwise.

   (b) Write a probability distribution that is uniform over all independent sets on a given graph (represented as elements of $\{0, 1\}^{|V|}$, and represent it as a factor graph in the example given above.

   To every edge $(ij) \in E$, define a function:

$$f_{(ij)}(\sigma_i^S, \sigma_j^S) = \mathbb{I}\left(\left(\sigma_i^S, \sigma_j^S\right) \neq (1,1)\right) = \begin{cases} 1 & \text{if} \quad \left(\sigma_i^S, \sigma_j^S\right) \neq (1,1) \\ 0 & \text{otherwise} \end{cases} . \tag{1}$$

   Or in words: $f_{(ij)}$ is one if at most one of the nodes i, j connected by the edge (ij) belong to S. With this definition and the bijection at point above, we can characterize an independent subset $S \subset V$ as:

$$S \text{ is independent} \quad \Leftrightarrow \quad \text{for all distinct } i, j \in S, \quad f_{(ij)}(\sigma_i, \sigma_j) = 1 \tag{2}$$

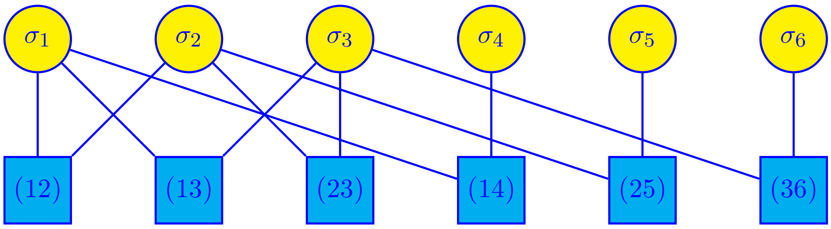
Or in words: an independent set is such that none of the nodes belonging to it are connected by an edge of the graph. For a given set of nodes $\vec{\sigma} \in \{0,1\}^N$ the uniform probability measure over independent sets is given by:

$$\mathbb{P}(\vec{\sigma}) = \frac{1}{\mathcal{Z}_G} \prod_{(ij)\in E} \mathbb{I}\left((\sigma_i, \sigma_j) \neq (1,1)\right) \tag{3}$$

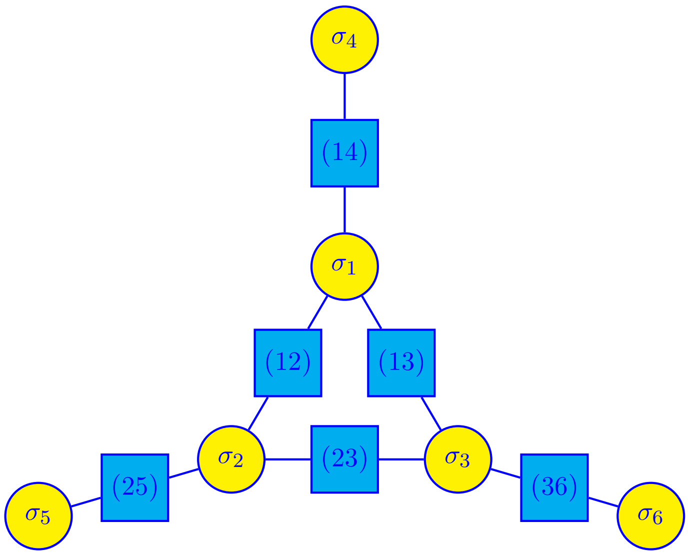
where:

$$\mathcal{Z}_G = \sum_{\vec{\sigma}\in\{0,1\}^N} \prod_{(ij)\in E} \mathbb{I}\left((\sigma_i, \sigma_j) \neq (1,1)\right) = \text{number of independent sets in } G \tag{4}$$

The factor graph associated with the example graph is:



Another way to represent the factor graph is



highlighting the fact that the factor graph is isomorphic to the original graph. This is not always the case, and in this example hinges on the fact that each factor node involves only two variable nodes, so that the factor can really be thought of as a weight associated to the edge.

(c) Write a probability distribution that gives a larger weight to larger independent sets, where the size of an independent set is simply its cardinality $|S|$. Represent it as a factor graph for the example given above.
Hint: many probability distributions assign more weight to $|S|$, but some choices lead to simpler factor graphs…

Note that the size of a set $|S|$ can be expressed in terms of the variables σ_i^S as:

$$|S| = \sum_{i=1}^{N} \sigma_i^S \tag{5}$$

To assign a larger weight to independent sets which are larger, we just need to multiply our density by any positive increasing function $g(|S|)$:

$$\mathbb{P}(\vec{\sigma}) = \frac{1}{\tilde{\mathcal{Z}}_G} g\left(\sum_{i=1}^{N} \sigma_i\right) \prod_{(ij)\in E} \mathbb{I}\left((\sigma_i, \sigma_j) \neq (1,1)\right) \tag{6}$$

For example, we can choose $g(x) = e^{hx}$ for $h > 0$ to get:

$$\mathbb{P}(\vec{\sigma}) = \frac{1}{\tilde{\mathcal{Z}}_G} \prod_{i=1}^{N} e^{h\sigma_i} \prod_{(ij)\in E} \mathbb{I}\left((\sigma_i, \sigma_j) \neq (1,1)\right) \tag{7}$$

Note that this would introduce a local factor node to variable node in the factor graph. Other choices of g are possible, but they may not nicely factorize. For this choice of g, the factor graph is identical as before, with a single-variable additional factor for each variable with weight $e^{h\sigma_i}$.

(d) Write the Belief Propagation equations for these problems (without coding or solving them) and the expression for the Bethe free energy that would be computed from the BP fixed points.

We shall write the BP equations in the case (7) of question (b). Note that the BP equations for the case discussed in (a) can be readily recovered simply by setting the field h to zero. We have

$$g_i(\sigma_i) = e^{h\sigma_i}, \qquad f_{(ij)}(\sigma_i, \sigma_j) = \mathbb{I}\left((\sigma_i, \sigma_j) \neq (1,1)\right). \tag{8}$$

This gives us the messages

$$\psi_{\sigma_i}^{(ij)\to i} = \frac{1}{Z^{(ij)\to i}} \left[\chi_1^{j\to(ij)} \delta_{\sigma_i,0} + \chi_0^{j\to(ij)}\right], \tag{9}$$

$$\chi_{\sigma_i}^{i\to(ij)} = \frac{1}{Z^{i\to(ij)}} e^{h\sigma_i} \prod_{(ik)\in E\setminus(ij)} \psi_{\sigma_i}^{(ik)\to i}. \tag{10}$$

The terms of the Bethe free entropy read

$$Z^i = \sum_{\sigma} e^{h\sigma} \prod_{j\in\partial i} \psi_{\sigma}^{(ij)\to i}, \tag{11}$$

$$Z^{(ij)} = \chi_0^{i\to(ij)}\chi_0^{j\to(ij)} + \chi_0^{i\to(ij)}\chi_1^{j\to(ij)} + \chi_1^{i\to(ij)}\chi_0^{j\to(ij)}, \tag{12}$$

$$Z^{i,(ij)} = \sum_{\sigma} \chi_{\sigma}^{i\to(ij)} \psi_{\sigma}^{(ij)\to i}. \tag{13}$$

2. **Matching problem**

The matching problem is another classical problem of graph theory. It is related to the dimer problem in statistical physics, where you aim at covering a graph with two-site dimers. Given a (unweighted, undirected) graph $G(V, E)$ a matching $M \subseteq E$ is defined as a subset of edges such that if $(ij) \in M$ then no other edge that contains node i or j can be in M. In other words a matching is a subset of edges such that no two edges of the set share a node.

Example problem: same as the independent graph one.

(a) Write a probability distribution that is uniform over all matchings on a given graph, and draw the factor graph corresponding to the example graph given for the independent set problem. Hint: again, find a Boolean encoding for a matching, similarly as what we did for the independent set.

The construction of the factor graph for matching is very similar to the one for the independent set, with the crucial difference that the variable nodes are the edges of G, instead of the nodes. As before, we start by assigning a binary variable to each edge of G which identifies whether it belongs or not to M:

$$s_{(ij)} = \begin{cases} 1 & \text{if } (ij) \in M \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

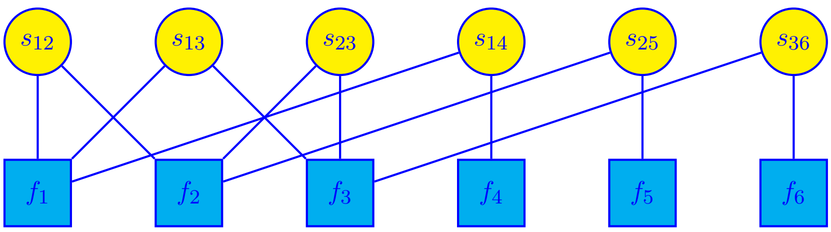
Let $N = |V|$. As before, for every node $i = 1, \cdots, N$ we assign a function which is zero if the node is attached to two edges belonging to M:

$$f_i\left(\{s_{(ij)}\}_{j\in\partial i}\right) = \mathbb{I}\left(\sum_{j\in\partial i} s_{(ij)} \leq 1\right) \tag{15}$$

Note that with this definition we allow for nodes to be unpaired. If we would like only perfect matchings (i.e. when all eges are paired), we would impose equality. The uniform measure over all matchings in G can therefore be written as:

$$\mathbb{P}(\vec{s}) = \frac{1}{\mathcal{Z}_G} \prod_{i=1}^{N} \mathbb{I}\left(\sum_{j\in\partial i} s_{(ij)} \leq 1\right) \tag{16}$$

where the partition function $\mathcal{Z}_G$ counts the total number of matching sets M in G. Note that different from (2), $\vec{s} \in \mathbb{R}^{|E|}$. Note that the uniform measure assigns the same weight to large matchings (i.e. when as many edges as possible are matched) and smaller matchings (e.g. when only half of the edges are matched) . To illustrate the factor graph of the matching problem, consider the same graph as in problem (2). The associated factor graph is given by:



Notice that in this case the factor graph is not isomorphic to the original graph, due to multibody factors that involve more than 2 variable nodes.

(b) Write a probability distribution that that gives a larger weight to larger matchings, where the size of a matching is the cardinality $|M|$. Then, draw the factor graph corresponding to the example graph given for the independent set problem

As before, we can write the size of a matching set as a function of $\vec{s}$:

$$|M| = \sum_{(ij)\in E} s_{(ij)} \tag{17}$$

Therefore, to assign a bigger weight to larger matchings, we just need to multiply the measure by any positive increasing function $g(|M|)$:

$$\mathbb{P}(\vec{s}) = \frac{1}{\mathcal{Z}_G} g\left(\sum_{(ij)\in E} s_{(ij)}\right) \prod_{i=1}^{N} \mathbb{I}\left(\sum_{j\in\partial i} s_{(ij)} \leq 1\right) \tag{18}$$

This is the softer way to encourage a perfect matching than to impose equality at the factor function f_i.

(c) Write the Belief Propagation equations for these problems (without coding or solving them) and the expression for the Bethe free energy that would be computed from the BP fixed points.

We treat the general case (b). Setting $g(\cdot) = 1$ allows to recover the usual matching problem. For the problem to be tractable, we have to assume that the bias $g(\cdot)$ factorizes, so we take

$$g\left(\sum_{(ij)\in E} s_{(ij)}\right) = e^{h \sum_{(ij)\in E} s_{(ij)}}. \tag{19}$$

So to connect with the lecture notes,

$$g_{(ij)} = e^{hs_{(ij)}}, \qquad f_i(\{s_{(ij)}\}_{j\in\partial_i}) = \mathbb{I}\left(\sum_{j\in\partial i} s_{(ij)} \leq 1\right). \tag{20}$$

The messages are then given by

$$\psi^{i\to(ij)}_{s_{(ij)}} = \sum_{\{s_{(ik)}\}_{k\in\partial_i\setminus j}} \prod_{k\in\partial_i\setminus j} \chi^{(ki)\to i}_{s_{(ik)}} \mathbb{I}\left(\sum_{l\in\partial i} s_{(il)} \leq 1\right), \tag{21}$$

$$\chi^{(ij)\to i}_{s_{(ij)}} = e^{hs_{(ij)}} \psi^{j\to(ij)}_{s_{(ij)}}. \tag{22}$$

The three contributions for the Bethe free energy are

$$Z^{(ij)} = \sum_{s} e^{\beta s} \psi^{i\to(ij)}_{s_{(ij)}} \psi^{j\to(ij)}_{s_{(ij)}}, \tag{23}$$

$$Z^{i} = \sum_{\{s_{(ij)}\}_{j\in\partial_i}} \mathbb{I}\left(\sum_{j\in\partial i} s_{(ij)} \leq 1\right) \prod_{j\in\partial_i} \chi^{(ij)\to i}_{s_{(ij)}} \tag{24}$$

$$Z^{(ij),i} = \sum_{s} \psi^{i\to(ij)}_{s} \chi^{(ij)\to i}_{s} \tag{25}$$

10.2 The Ising model on d-regular random graphs

Consider the following probability distribution

$$p_G(S) = \frac{1}{Z} \exp\left(\beta \sum_{(ij)\in E} S_i S_j\right) \tag{26}$$

where G is a graph with N nodes, and edge set E. This is the Ising model on a graph G. Assume that G is a uniformly sampled d-regular graph, i.e. a graph whose nodes have all $d = O(1)$ neighbours.

1. Sketch the associated factor graph, and write the BP equations.

   The factor graph is the same as the one for the independent set problem, with one factor node with weight $e^{\beta S_i S_j}$ for each edge $(ij) \in E$. The BP equations are then

$$\psi_{\sigma_i}^{(ij)\to i} \propto \sum_{s=\pm 1} e^{\beta \sigma_i s} \chi_s^{j\to(ij)}, \tag{27}$$

$$\chi_{\sigma_i}^{i\to(ij)} \propto \prod_{(ik)\in E\setminus(ij)} \psi_{\sigma_i}^{(ik)\to i}. \tag{28}$$

2. When studying problems on random d-regular graphs, we often make a sort of RS ansatz by saying that the BP messages will be uniform over all edges of the graph. This is because all local neighbourhoods on the factor graphs are identical. Similarly as to an RS ansatz, this is not necessarily the only, or the correct solution to the BP equations, but for sure it is the simplest to study and worth looking at. Use this uniformity assumption to derive the following reduced BP equation

$$\chi(s) = \frac{\left[\sum_t e^{\beta st}\chi(t)\right]^{d-1}}{\sum_{s'}\left[\sum_t e^{\beta s' t}\chi(t)\right]^{d-1}} \tag{29}$$

   where for all nodes i and edges (ij) we called $\chi = \chi^{i\to(ij)}$.

   If all messages are the same, the BP equations reduce to

$$\chi(s) = \frac{\psi(s)^{d-1}}{\sum_{s'}\psi(s')^{d-1}}, \qquad \psi(s) = \frac{\sum_t e^{\beta st}\chi(t)}{\sum_{s',t} e^{\beta s' t}\chi(t)} \tag{30}$$

   which can be closed into a single equation for χ, leading to the result.

3. Compute the marginal over a single spin s under the uniform ansatz.

   We have

$$\mu(s) = \psi(s)^d = \frac{\left[\sum_t e^{\beta st}\chi(t)\right]^{d}}{\sum_{s'}\left[\sum_t e^{\beta s' t}\chi(t)\right]^{d}} \tag{31}$$

4. Show that the paramagnetic fixed point $\chi(s) = 1/2$ is a solution of the BP equations for all β. Why do we call this paramagnetic?

Substituting the ansatz into the BP equation gives

$$\chi(+1) = \frac{1}{2} = \frac{\left[\sum_t e^{\beta t}\right]^{d-1}}{\sum_{s'}\left[\sum_t e^{\beta s' t}\right]^{d-1}} = \frac{\left[e^{\beta} + e^{-\beta}\right]^{d-1}}{\left[e^{\beta} + e^{-\beta}\right]^{d-1} + \left[e^{-\beta} + e^{\beta}\right]^{d-1}} \tag{32}$$

which is satisfied. It is a paramagnetic fixed point because the marginal over a spin becomes uniform over $\{-1, +1\}$, implying average zero magnetization.

5. Show that the ferromagnetic fixed point $\chi(+1) = a \in [0,1]$ and $\chi(-1) = 1 - a$ is a solution of the BP equations for some value of a_*, and show that a_* satisfies a scalar equation.

Substituting the ansatz into the BP equation gives

$$\chi(+1) = a = \frac{\left[ae^{\beta} + (1-a)e^{-\beta}\right]^{d-1}}{\left[ae^{\beta} + (1-a)e^{-\beta}\right]^{d-1} + \left[ae^{-\beta} + (1-a)e^{\beta}\right]^{d-1}} = \left[1 + \left(\frac{ae^{-\beta} + (1-a)e^{\beta}}{ae^{\beta} + (1-a)e^{-\beta}}\right)^{d-1}\right]^{-1} \tag{33}$$

which is the equation that a_* needs to satisfy.

6. Compute the average magnetization as a function of a^*.

We have

$$m = \langle \frac{1}{N}\sum_{i=1}^{N} s_i \rangle = \mathbb{E}_{s\sim\mu}[s] = 2a^* - 1 \tag{34}$$

. We thus see that a is a proxy for the magnetization.

7. We expect that in this model there is a second order phase transition, as this is the case for $d \to \infty$ (Curie-Weiss model), as well as for the finite dimensional counterparts of the Ising model. To derive the second order phase transition threshold β_c we can study the stability of the iteration

$$a_{t+1} = \left[1 + \left(\frac{a_t e^{-\beta} + (1-a_t)e^{\beta}}{a_t e^{\beta} + (1-a_t)e^{-\beta}}\right)^{d-1}\right]^{-1} = f(a_t) \tag{35}$$

around the paramagnetic solution $a = 1/2$. We expect that the iteration will fall back on the paramagnetic solution in the paramagnetic phase, while it will diverge away from it in the ferromagnetic phase. In other words, $a = f(a)$ is our state equation, and we are checking whether $a^* = 1/2$ is a maximum of the associated free entropy.
Argue that the iteration initialized at $a_0 = 1/2 + \epsilon$ for small ϵ converges back to the paramagnetic solution only if $f'(1/2) < 1$.

If $a_t = 1/2 + \epsilon$ then

$$a_{t+1} - 1/2 = f(1/2+\epsilon) - 1/2 = f(1/2) + \epsilon f'(1/2) + \dots - 1/2 = \epsilon f'(1/2) + \dots \tag{36}$$

from which we see that if $f'(1/2) < 1$, the iterates will get closer to $1/2$ if $f' < 1$, and farther viceversa.

8. Compute the critical threshold $\beta_c(d)$ as a function of the degree d.

We just need to compute $f'(1/2) = 1$ and solve for β. We have

$$f(a) = \text{logistic}\left((1-d)\log\left(\frac{a_t e^{-\beta} + (1-a_t)e^{\beta}}{a_t e^{\beta} + (1-a_t)e^{-\beta}}\right)\right) \tag{37}$$

where $\text{logistic}(x) = (1+e^{-x})^{-1}$. Computing the derivative gives

$$f'(1/2) = (d-1)\tanh(\beta) = 1 \implies \beta_c(d) = \text{arctanh}\left(\frac{1}{d-1}\right) \tag{38}$$

9. Compare the value of $\beta_c(d)$ with the value for the phase transition of the 1d Ising model $\beta_c = +\infty$, and with the value for the Curie-Weiss model $\tilde{\beta}_c = 1$ (notice that there is a difference in normalization between this Ising model and the Curie-Weiss model!). To which values of d the two correspond? How should we rescale β in our problem to be in the same scaling as the Curie-Weiss model?

   For $d = 2$ we reduce to the 1d chain model, and $\beta_c = +\infty$. To compare with the Curie-Weiss, we need to make sure that the Hamiltonian $-\sum_{(ij)\in E} S_i S_j$ is of order $O(N)$ when $d = N-1$. But if $d \gg 1$, the Hamiltonian is of order $O(dN)$, so we want to divide it by $1/d$. Thus, calling $\beta = \tilde{\beta}/d$ we obtain

$$\tilde{\beta}_c(d) = d\beta_c(d) = d\,\text{arctanh}\left(\frac{1}{d-1}\right) \to 1 \tag{39}$$

   as $d \to \infty$.