Stability-capacity diagram of a neural network with Ising bonds
E. Amaldi, S. Nicolis

To cite this version:

HAL Id: jpa-00211064
https://hal.archives-ouvertes.fr/jpa-00211064
Submitted on 1 Jan 1989

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Stability-capacity diagram of a neural network with Ising bonds

E. Amaldi (1, *) and S. Nicolis (2)

(1) Caltech Concurrent Computation Program, California Institute of Technology, Pasadena, CA 91125, U.S.A.
(2) Dipartimento di Fisica, Università di Roma, « La Sapienza », piazzale Aldo Moro 2, Roma 00185, Italy

(Reçu le 3 mars 1989, accepté sous forme définitive le 16 mai 1989)

Abstract. — The stability-capacity diagram of a fully connected neural network, whose bonds take the values 1 and −1 with equal probability, is determined numerically. Two different optimization methods (simulated annealing and tabu search) are used and their relevant features are discussed. The results indicate the existence of a region, in the stability-capacity plane, where replica-symmetry should be broken, within the replica-symmetric phase found by previous, analytical, computations. The critical capacity, consistent with simple arguments on information storage, is found to be less than one.

1. Introduction.

A neural network is composed of many elementary units (called, quite naturally, « neurons »), that are assumed to be in one of two possible states: active or inactive. This is usually quantified by a function $S$, from the set of neurons to the set $\{1, -1\}$, such that $S_i = 1$ if neuron $i$ is active and $S_i = -1$ if it is inactive.

It may happen that the state of neuron $i$ has some influence on whether neuron $j$ is active or not. The simplest way is to assume that the two neurons are directly connected by

(*) Work done mainly during a stay at the University of Rome. The completion of this research was supported by DOE grant DE-FG-03-85ER25009 and the Program Manager of the Joint Tactical Fusion Office.
a synaptic bond of strength $J_{ij}$. Note that the bond may be asymmetric; $J_{ij} \neq J_{ji}$. The quantitative expression of this « influence » is assumed to be the following: if neuron 1 is connected to neurons 2, ..., $z$, that are in states $S_2, ..., S_z$, by bonds of synaptic strength $J_{12}, ..., J_{1z}$, the state of neuron 1 will be determined by

$$S_1 = \text{sign}\left(\sum_{k=2}^{z} J_{1k} S_k - \kappa_1 \sqrt{z} S_1\right)$$

(1)

where $\text{sign}(x) = 1$ if $x \geq 0$, $-1$ if $x < 0$ and $\kappa_1$ is the stability for neuron 1, introduced by Gardner [1]. This parameter will be taken the same for all neurons from now on. It is worth remembering, however, that it need not be. Another simplifying assumption, that will be in effect from now on, is that a synaptic bond exists between all pairs of neurons, i.e. the network is fully connected. The total number of neurons will be denoted by $N$. In the equation above, therefore, $z = N$ and the same equation (with the obvious changes) will hold for neurons 2, ..., $N$. Finally (as was tacitly assumed in the form of Eq. (1)), the neurons do not self-interact; there is no synaptic bond between neuron 1, say, and itself: $J_{ii} = 0$ for all neurons.

The state of the network will be defined as the $N$-component vector $(S_1, ..., S_N)$. Equation (1) amounts to a self-consistency condition on the components of this vector. Henceforth, the term « states » will refer to those vectors that satisfy this self-consistency condition.

The fundamental problem here is the following: consider $pN$-component vectors, with components 1 and $-1$. It is assumed that these vectors are statistically uncorrelated-introducing correlations turns out to complicate matters in a fairly non-trivial way and will not be considered here. Another assumption (not totally unrelated to the previous one, [2]) will be that the values 1 and $-1$ are chosen with equal probability. Is it possible to find a set of $J_{ij}$'s, so that any such $p$ vectors be states of the network (as defined above)? It is clear that, if a vector is a state of the network for some value of the stability, say $\kappa^{(1)}$, it will remain a state for all values less than $\kappa^{(1)}$. The problem may then be stated as the search for the greatest value of the stability such that there exists a set of $J_{ij}$'s for which $p$ vectors, satisfying the above assumptions, are states of the network. This problem has been solved when $J_{ij}$ and $J_{ji}$ are independent random variables that obey the so-called « spherical constraint »

$$\sum_{j=1}^{N} J_{ij}^2 = N$$

for all $i$ in [1, 3]. Another constraint was studied in reference [4], the so-called « Ising » constraint: $J_{ij}$ and $J_{ji}$ are independent random variables that take the values 1 and $-1$ with equal probability. The stability-capacity diagram was derived under the assumption of replica symmetry, which was found to be inconsistent, however. The purpose of this note is to determine this diagram numerically. In the process, we shall describe a new optimization technique, tabu search, which seems a promising tool for the numerical study of neural networks.


The strategy will be to fix $N$, vary $p$ and determine $\kappa$ as follows: generate the $J_{ij}$'s and $pN$-component vectors, $S_1^{\mu}, i = 1, ..., N, \mu = 1, ..., p$. Fix a site $i$ at random and calculate

$$h_i^{\mu} = S_i^{\mu} \left(\sum_{j \neq i}^{N} J_{ij} S_j^{\mu}\right)$$

(2)
Since the $J_{ij}$'s are independent random variables, the particular choice of $i$ doesn't matter, and will be omitted in the following. Find now the smallest $h^{\mu}$. This will be the candidate $\kappa$; the question now arises, whether it is possible to improve on this value.

This defines a typical combinatorial optimization problem: its parameters are the number, $N$, of neurons, that are pair-wise connected by bonds $J_{ij}$—these bonds take the values 1 and $-1$ with equal probability—and the number, $p$, of statistically uncorrelated $N$-component binary vectors. An instance is defined by assigning specific to $N$, $p$ and by a set of $p$ such vectors $\{S^{\mu}\}$, $\mu = 1, \ldots, p$ and a solution is a collection of $J_{ij}$'s that maximize

$$\kappa = \min_\mu \left\{ \frac{S^{\mu}_i}{\sqrt{N}} \sum_{j \neq i} J_{ij} S^{\mu}_j \right\}$$

for a given $i$ (and a given $p$). This is a non-trivial task because of the existence of many local extrema, i.e. configurations that yield values of $\kappa$, that decrease upon all possible single-bond-flips. To overcome this obstacle, we used two algorithms that can escape from most, if not all, such undesirable situations, simulated annealing and tabu search.

The « simulated annealing » algorithm, introduced in [5], proceeds as follows: bond-flips are carried out; that is, $J_{ij} \leftarrow -J_{ij}$. A new set of $h^{\mu}$'s is calculated and the bond-flip is accepted if the new smallest $h^{\mu}$ is greater than the previous one, otherwise it is rejected with probability $1 - \exp(-\Delta h/T)$, where $T$ is a « temperature » that is initialized at some suitably high value (on the scale of the $h$'s) and then reduced according to some annealing schedule and $\Delta h = h_{\text{new}} - h_{\text{old}}$ (where the smallest $h$'s of each set are taken). This continues until all bonds $J_{ij}$ are tried. The bond-flips are carried out many times at the same temperature. The final value of $h$, i.e. the value of the smallest element of the set $\{h^{\mu}\}$ at the last operation, is taken as the value of the stability for the values of $p$ and $N$ under consideration (it must be divided by $\sqrt{N}$, of course, for comparison with results obtained at different values of $N$). The value of $\alpha = p/N$, for which $\kappa = 0$, denoted $\alpha_c$, indicates the maximum number of vectors that may be states of a network, whose bonds satisfy the constraint under consideration (in this case the Ising constraint). We call $\alpha$ the capacity of the network.

The other method applied to this problem is the « tabu search » technique, introduced in [6], applied to some « hard » optimization problems in [7] and [8] and adapted to neural network problems in [9].

Tabu search works as follows: a given number, $n_b$, of bond-flips—chosen at random—is considered, at each step, and the corresponding values of $\kappa$ are calculated as before; to bond-flip $J_{i1} \leftarrow -J_{i1}$ corresponds a new value of $\kappa$, $\kappa^{(1)}_{\text{new}} = \min (h^1, \ldots, h^p)$, and so on; to bond-flip $J_{i\mu} \leftarrow -J_{i\mu}$, corresponds $\kappa^{(n_b)}_{\text{new}} = \min (h^1, \ldots, h^p)$. The largest of these « partial » values is then accepted as the new value of $\kappa$:

$$\kappa_{\text{new}} = \max (\kappa^{(1)}_{\text{new}}, \ldots, \kappa^{(n_b)}_{\text{new}}).$$

As will be seen at an arbitrary step, not all of the $n_b$ bond-flips may be allowed. It should be stressed that this value is accepted, whether it is greater than that found at the previous step, or not. This allow the algorithm to leave local maxima, but, obviously, implies the hazard of revisiting the same bond-configurations many times. To exclude bond-flips, that could lead to $J_{ij}$-configurations already tested, the bonds, that have been accepted during the last $L$ steps, are put on a « tabu » list $T$ and may not be flipped. This list is initially empty and updated in the following way:

the $\kappa_{\text{new}}$, that was defined above corresponded to a particular bond-flip, say $J_{ik} \leftarrow -J_{ik}$. This bond is now placed in the $L$th position of the tabu list. During the next
$L$ steps (rounds of $n_b$ trials), this bond will not be considered. The $\kappa$ found in the step immediately following, will correspond to bond-flip $J_{it} \leftarrow - J_{it}$. This bond will now occupy position $L$ of the tabu list and bond $J_{ik}$ will move to position $L - 1$. When a bond moves past position 1 of the list it may again be considered.

It may well happen, however, that, at some point during the search procedure, were it possible to flip a, momentarily, « tabooed » bond, one would obtain a better value for the stability than by flipping one of the « allowed » bonds. The tabu list would not change, of course, but the bond configuration would; and this might lead to better regions in search space. To this end, we shall define an « aspiration level function $A$ », on the set of $\kappa$ values. This function works in the following way (see also the explicit example of the Appendix):

Let us assume, for instance, that the procedure, outlined above, has already been carried out $M$ times. The value of $\kappa$ found during the $M$th step is called $\kappa_{\text{present}}$ and we shall assume, for simplicity, that it is the first time this value has been encountered. Initially we set $A(\kappa) = \kappa$ for every value of $\kappa$ and, in particular, for $\kappa_{\text{present}}$. The $\kappa_{\text{new}}$ found at the $M + 1$st step, is the largest value of $\kappa$ corresponding to one of the $n_b$ bond-flips not on the tabu list, say $J_{im} \leftarrow - J_{im}$. Similarly, $\kappa'_{\text{new}}$ denotes the largest $\kappa$ associated to one of the $n_b$ bond-flips that are tabooed, say $J_{it} \leftarrow - J_{it}$. Then, if

\[
\kappa'_{\text{new}} > \kappa_{\text{new}}
\]

and

\[
\kappa'_{\text{new}} > A(\kappa_{\text{present}}),
\]

we accept the bond-flip $J_{it} \leftarrow - J_{it}$ and not the bond-flip $J_{im} \leftarrow - J_{im}$; the tabooed bond-flip has overridden its taboo status but remains, of course, on the tabu list. Furthermore, the function $A$ is updated in the following way

\[
A(\kappa_{\text{present}}) \leftarrow \kappa'_{\text{new}}.
\]

In a contrary case we accept the « allowed » bond-flip, which enters the tabu list at position $L$. If $\kappa_{\text{new}} > A(\kappa_{\text{present}})$, moreover, then $A(\kappa_{\text{present}}) \leftarrow \kappa_{\text{new}}$. This means that if, after a future step, we find the same value of $\kappa_{\text{present}}$, a « tabooed » bond-flip must lead to a value at least greater than $\kappa_{\text{new}}$ (or $\kappa'_{\text{new}}$ as the case may be) in order to override its tabu status.

The search is stopped after $K$ steps, where $K$ is arbitrarily chosen. A worked example is discussed in detail in the appendix.

3. Results. Discussion.

Both procedures were carried out for $N = 27, 45, 63$ and 81. The choice of $N$ odd is a technicality so that $h^\alpha$, which is thus the sum of an even number of terms, that take the values 1 or $-1$, may take the value zero. The annealing schedule

\[
T_n = 5.0(0.95)^n, \quad n = 0, \ldots, 99
\]

was found by trial and error. It leads to freezing, at $T \leq 0.13$, for all values of $N$ and $p$, which provides a convenient halting point, and changing the prefactor does not lead to higher values of $\kappa$. A slower schedule, $T_n = 5.0(0.98)^n, n = 0, \ldots, 181$, was also tried in several cases, but did not give new results. Of course this does not preclude the existence of an optimal annealing schedule, different from these. A slower schedule, however, soon becomes prohibitive, in terms of CPU time, for $N \geq 81$; it may well run for 3 hours and more (per
given $p$ and per configuration) on a VAX 8650. Every complete bond-flip operation was carried out 200 times at the same temperature; the first 50 steps were discarded during the averaging. Increasing the number of steps to 350 did not change the results.

In tabu search $n_t$ was taken equal to $N - 1$; that is, all bonds were tried. In this case the algorithm is deterministic. In accord with empirical evidence \cite{7}, $L = 7$ was found to give satisfactory results. $L = 14$ was also tried and did not lead to significant changes; the same was true for $7 \leq L \leq N/2$ (for $L \leq 7$ we run the risk of going around in circles about some configuration; for $L$ too large of getting stuck with no good and legal bond-flips available).

The choice of $K$ is much more complex—it, probably, also depends upon $L$. It is here that a fundamental difference between tabu search and simulated annealing appears. We recall that simulated annealing picked as value for $K$ the smallest among the $\{h^\mu\}$'s calculated at the last iteration. It is possible (though—in the cases studied here—rare) that a greater value for $K$ appear at some earlier step. Since freezing occurs, the probability of a «fluctuation» of appreciable size is low.

In the case of tabu search the situation is not so clear. Freezing, in the thermodynamic interpretation associated with simulated annealing, does not occur. The values of $K$ fluctuate from iteration to iteration and we pick, as final $K$, the largest value of \( \min \{ h^\mu \} \) found till the current iteration step. It is thus possible that increasing $K$ would lead to higher values for $K$, with no indication of a «natural» halting point of the procedure. In practice, fortunately, this was not the case. If one would plot $K$ vs. iteration step, one would find a single, fairly pronounced, maximum; after some iterations the fluctuations remain strong but not enough to lead to a significantly (see below) larger value of $K$. Various values of $K$ were tried; $K = 6$.000 was found satisfactory, which means that the largest value of $K$ occurred well before, in most cases, and when it did occur quite close to the end, a run at $K = 18$.000 did not lead, in most of those cases, to new results; when it did the change was not significant (e.g. it amounted to a change of 2, that is, a bond-flip). This remained true for $K$ up to 7.000 as well. Choosing $L$ and $K$ is, at the moment, as much a question of trial and error as the determination of the annealing schedule, so slight (or not so slight) improvements may be possible through a judicious choice of their values.

It might be useful to try to understand the difference between the «freezing» in simulated annealing and the behavior of tabu. In the context of simulated annealing freezing means that all bond-flips tend to decrease $K$, which implies a local extremum; also that the bond-flips lead to changes of $\Delta h$ so great, with respect to the current low value of the temperature, that they are rejected with probability very close to 1.

In tabu, on the other hand, since we always change one bond, the value of $K$ will always change, which leads to the fluctuations mentioned earlier; these will, of course, also depend on the existence of the tabu list as well as on its length $L$.

Both methods gave consistent values for $K$ (see Tab. I)—though tabu search results were consistently higher than simulated annealing. It is probable that a much slower annealing schedule (especially for large $N$ and $\alpha$ close to 1) would yield higher values for $K$; unfortunately, simulated annealing becomes prohibitively time-consuming here.

Tabu search was also much faster. This may be traced to the fact that simulated annealing involved the call of a random number generator and a logical IF for every update attempt, whereas tabu search did not. As is well known, these operations are particularly time-consuming in problems of this type (this is no longer true for more complex objective functions). It seems possible to optimize these operations, however, so the difference between the running times may decrease considerably. In most of the tabu search cases the final value for $K$ was found in less than 1.000 iterations and the procedure was about 20 times faster than simulated annealing.
Table I. — Comparison between tabu search (upper row) and simulated annealing (lower row) for the stability. To every entry must be assigned a systematic error $\pm 1/\sqrt{N}$. $N$ is the number of neurons and $\alpha$ the ratio of stored patterns to number of neurons.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>27</th>
<th>45</th>
<th>63</th>
<th>81</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.11</td>
<td>2.013</td>
<td>1.949</td>
<td>1.996</td>
<td>1.949</td>
</tr>
<tr>
<td></td>
<td>2.013</td>
<td>1.944</td>
<td>1.938</td>
<td>1.844</td>
</tr>
<tr>
<td>0.22</td>
<td>1.125</td>
<td>1.124</td>
<td>1.144</td>
<td>1.095</td>
</tr>
<tr>
<td></td>
<td>1.125</td>
<td>1.101</td>
<td>1.085</td>
<td>1.111</td>
</tr>
<tr>
<td>0.33</td>
<td>0.711</td>
<td>0.711</td>
<td>0.659</td>
<td>0.650</td>
</tr>
<tr>
<td></td>
<td>0.681</td>
<td>0.711</td>
<td>0.659</td>
<td>0.615</td>
</tr>
<tr>
<td>0.44</td>
<td>0.444</td>
<td>0.436</td>
<td>0.368</td>
<td>0.273</td>
</tr>
<tr>
<td></td>
<td>0.444</td>
<td>0.413</td>
<td>0.329</td>
<td>0.173</td>
</tr>
<tr>
<td>0.55</td>
<td>0.355</td>
<td>0.206</td>
<td>0.078</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>0.266</td>
<td>0.138</td>
<td>0.097</td>
<td>0.017</td>
</tr>
<tr>
<td>0.66</td>
<td>0.178</td>
<td>0.069</td>
<td>0.019</td>
<td>$-0.034$</td>
</tr>
<tr>
<td></td>
<td>0.088</td>
<td>0.069</td>
<td>$-0.058$</td>
<td>$-0.256$</td>
</tr>
<tr>
<td>0.77</td>
<td>0.000</td>
<td>$-0.046$</td>
<td>$-0.155$</td>
<td>$-0.205$</td>
</tr>
<tr>
<td></td>
<td>$-0.059$</td>
<td>$-0.137$</td>
<td>$-0.174$</td>
<td>$-0.376$</td>
</tr>
<tr>
<td>0.88</td>
<td>$-0.030$</td>
<td>$-0.161$</td>
<td>$-0.236$</td>
<td>$-0.325$</td>
</tr>
<tr>
<td></td>
<td>$-0.118$</td>
<td>$-0.252$</td>
<td>$-0.271$</td>
<td>$-0.496$</td>
</tr>
<tr>
<td>1.00</td>
<td>$-0.266$</td>
<td>$-0.252$</td>
<td>$-0.271$</td>
<td>$-0.393$</td>
</tr>
<tr>
<td></td>
<td>$-0.385$</td>
<td>$-0.298$</td>
<td>$-0.368$</td>
<td>$-0.632$</td>
</tr>
</tbody>
</table>

The stability-capacity diagram is displayed in figure 1, in which the theoretical curve, derived in [4] under the assumption of replica symmetry, is also included. The error bars are from statistical errors (13 configurations per point) and systematic (that are taken $\pm 1/\sqrt{N}$ since $\kappa \sqrt{N} = h$ is an even integer — the least difference between two configurations is a bond-flip which corresponds to a difference in $h$ of two). Specifically, we take as error bar max (statistical, systematic). We have found that the systematic error was always greater than the statistical one.

Another issue pertains to the relevance of these simulations towards understanding the limit $N \to \infty$. From the figures it seems that, as $N$ increases, the curves do tend to « bunch » together, signalling that a saturation limit may have been reached. This may be taken as evidence that, indeed, the results obtained on the samples studied here do not change qualitatively with increasing system size. This impression is strengthened by figure 2, where we have attempted to plot the stability $\kappa$ vs. $1/N^2$, $x = 0.3$, in the range $\alpha = 0.2 - 0.4$. This is the range where $\kappa$ varies most rapidly as a function of the capacity $\alpha$, which renders interpolation easier (see Fig. 1). The extrapolation to $1/N = 0$ seems quite hazardous; plotting the obtained values on the same figure as the other points shows, however, that the approximation is not that bad, after all. Special attention, in this regard, should be paid to a feature, particular to this model, namely, the « discreteness » of the diagram is due to the fact that $\kappa N^{1/2}$ may take only integer values. The true curve, must therefore, interpolate, taking this into account. This fact highlights even more the difficulty of reliably estimating $\alpha_c$.

There are several problems associated with computing the stability-capacity diagram at high values of $\alpha$ (and, incidentally, reliably estimating the critical capacity, $\alpha_c$): the curve becomes very flat (especially beyond 0.4); the systematic error of one bond-flip, (in the
vertical direction) leads to a much larger error in the horizontal direction. We may, therefore, only establish bounds on \( \alpha_c \). Another point is related to the fact that the stability values are determined up to a bond-flip; near the critical capacity this leads to fluctuations that may trap the system in metastable states with negative (or positive) stability, getting out of which may take a long time.

A most noteworthy feature of the figures is the following: we recall that, in the region below the curve, replica-symmetry is stable [4]. Our curve, on the other hand, always lies below that calculated analytically. This means that the replica-symmetric region is, in fact, quite smaller than expected. It would be of much interest to study how replica-symmetry is broken in the « intermediate » region. A standard tool is the overlap distribution [10]. As regards analytical approaches to replica-symmetry breaking [10], the existence of the « intermediate region », seems to indicate that new ideas might be needed.

Finally, a remark concerning the critical capacity \( \alpha_c \) is in order. The replica-symmetric computation [4] gave \( \alpha_c = 4/\pi \approx 1.26 \). This, by itself, suggested that replica-symmetry could not hold, since, by a simple argument, a neural network with \( O(N^2) \) binary bonds, could not store more than \( N \) \( N \)-component binary vectors, so \( \alpha_c \) must be less than one. Indeed our results indicate that \( \alpha_c \) lies between 0.6 and 0.9 \(^{(1)}\). In the absence of analytical computations, an accurate numerical determination of \( \alpha_c \) is still of some interest, though the difficulties should not be underestimated.

\(^{(1)}\) We recently received a preprint by Gardner and Derrida [11], where they calculate \( \alpha_c \) for small values of \( N \) (up to 15); extrapolating they predict \( \alpha_c = 0.75 \). Finite-size effects are extremely significant, however, as our results show.
Fig. 2. — Stability vs. $1/N^x$, $x = 0.3$; $\alpha = 0.2 - 0.4$. (Line obtained from least squares fit).
Fig. 2 (continued).

(d) \[ \alpha = 0.32 \]

(e) \[ \alpha = 0.36 \]

(f) \[ \alpha = 0.40 \]
Acknowledgements.

It is a pleasure to thank M. Virasoro for suggesting this problem, stimulating discussions and critical reading of the manuscript. We acknowledge useful comments from G. Fox. E. A. also wishes to thank D. de Werra and A. Hertz for introducing him to the field and to the tabu search method. In addition, he expresses his gratitude to the Physics Department of the University of Rome, « La Sapienza », for its hospitality. S. N. would also like to acknowledge helpful conversations with D. J. Amit, W. Krauth, M. Mézard and N. Sourlas.

Appendix

Worked example for tabu search.

In this appendix we wish to present in detail a simple example of how tabu search actually works. Despite its simplicity, this « toy » calculation will turn out to involve all the concepts discussed and will, hopefully, clarify them. We shall compare the result obtained by this method to the solution found by complete enumeration.

Consider the system, shown in figure 3, composed of $N = 5$ neurons. The fixed site is $i = 1$ so the bonds are $J_{ij}$, $j = 2, ..., 5$. We consider two vectors,

$$S^{(1)} = (1\ 1\ 1\ 1\ 1)$$
$$S^{(2)} = (1\ -1\ 1\ -1\ 1)$$  \hspace{1cm} (4)

so $p = 2$ and $\alpha = p/N = 0.4$. The problem is to calculate the greatest value of the stability $\kappa$, defined by (we omit the $1/\sqrt{N}$ factor, for the moment, and will put it back at the end)

$$\kappa = \min (h^1, h^2)$$  \hspace{1cm} (5)

$$h^1 = \sum_{j=2}^{5} J_{ij} S^{(1)}_j$$

$$h^2 = \sum_{j=2}^{5} J_{ij} S^{(2)}_j$$  \hspace{1cm} (6)

We generate an initial configuration of the bonds $J_{1j}$, say

$$J_1 = (1, -1, -1, 1) .$$

Fig. 3. — A toy model for tabu search.
This configuration leads to $\kappa = 0$. Does there exist a configuration that could give a higher value to $\kappa$? To answer this question, we search in the space of bonds for a « better » configuration by performing bond-flips $(J_{ij} \leftarrow - J_{ij}, j = 2, \ldots, 5)$ on the initial one. Following the notation introduced above, we shall perform $n_b = N - 1$ bond-flips per iteration, for, say, $K = 3$ iterations. Every time we accept a bond-flip, we shall put it « on quarantine » on a tabu list of length $L = 2$.

1. Iteration. $J_{12} \leftarrow - J_{12}$ gives $\kappa^2 = \min (2, -2) = -2$

$$J_{13} \leftarrow - J_{13}, \kappa^3 = \min (2, 2) = 2$$
$$J_{14} \leftarrow - J_{14}, \kappa^4 = \min (2, -2) = -2$$
$$J_{15} \leftarrow - J_{15}, \kappa^5 = \min (2, -2) = -2.$$

This gives

$$\kappa_{\text{new}}^{(1)} = \max (-2, 2, -2, -2) = 2$$

corresponding to the bond-flip $J_{13} \leftarrow - J_{13}$. The new configuration is $J_1 = (1, 1, -1, 1)$ and the tabu list is of the form

$$T = (J_{13}, \times) ;$$

bond $J_{13}$ is in position 2 and position 1 is, for the moment, empty. Since initially we have $\kappa_{\text{present}} = 0$ and $\kappa_{\text{new}}^{(1)} = A(0) = 0$, the aspiration function $A$ is updated accordingly: $A(0) = \kappa_{\text{new}}^{(1)} = 2$. This means that if, at a future step, we find $\kappa_{\text{present}} = 0$, at the step immediately following, a tabooed bond-flip must lead to $\kappa_{\text{new}}^{(1)} > 2$ and not zero, in order to override its quarantine.

2. Iteration. We repeat the process on the new configuration, omitting the bond-flip $J_{13} \leftarrow - J_{13}$. We obtain

$$\kappa_{\text{new}}^{(2)} = \max (0, 0, 0) = 0$$

but before we accept the corresponding bond-flip, $J_{12} \leftarrow - J_{12}$, we must evaluate the aspiration function and consider flipping the tabu bonds (here only $J_{13}$). This is a sticky point, so let’s do it slowly. Following the notation in the body of the paper, we have $\kappa_{\text{present}} = \kappa_{\text{new}}^{(1)} = 2$, so $A(2) = 2$. The idea is to see whether flipping $J_{13}$ we obtain a $\kappa_{\text{new}}^{(2)}$ strictly greater than this and than $\kappa_{\text{new}}^{(2)}$. If this is the case, we accept the bond-flip $J_{13} \leftarrow - J_{13}$ and not the bond-flip $J_{12} \leftarrow - J_{12}$ and the tabu list will remain unchanged. We find $\kappa_{\text{new}}^{(1)} = 0 < A(2)$, which means that $J_{13}$ can not be flipped. Since $A(\kappa_{\text{present}}) = \kappa_{\text{new}}^{(2)}$, we do not update the aspiration function either. The new configuration is $J_1 = (-1, 1, -1, 1)$, the tabu list becomes

$$T = (J_{12}, J_{13})$$

and $\kappa_{\text{present}} = 0$.

3. Iteration. We are left with only $J_{14}$ and $J_{15}$ to flip. We find $\kappa_{\text{new}}^{(3)} = \max (2, -2) = 2$ and tend to accept as new configuration $J_1 = (-1, 1, 1, 1)$. Let’s check the aspiration function. Since $\kappa_{\text{present}} = 0$, $A(0) = 2$; flipping $J_{12}$ gives $\kappa = 2$; flipping $J_{13}$ also gives $\kappa = 2$, so $\kappa_{\text{new}}^{(1)} = \max (2, 2) = A(0)$; and neither of the tabooed bonds may override its « quarantine ». The tabu list becomes

$$T = (J_{14}, J_{12})$$

and the aspiration function is not updated: $A(0)$ remains equal to 2.
The way tabu works is, hopefully, clear.

To conclude, tabu search leads to $\kappa = 2$ for this example. Since there are $2^d = 16$ possible configurations for the bonds $J_1$, we may also search for the maximum $\kappa$ by direct enumeration. We find that $\kappa$ may take four possible values $2, 0, -2$ and that there correspond two configurations to the value $\kappa = 2$, which is, in fact, the global maximum.

It is, of course, obvious that we should examine all pairs of vectors $S^{(1)}$ and $S^{(2)}$, find the greatest value for $\kappa$ in each case and take the average.

References