Landscape of solutions in constraint satisfaction problems

Marc Mézard,1 Matteo Palassini,1,2 and Olivier Rivoire1

1Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris-Sud, F–91405 Orsay, France.
2Departament de Física Fonamental, Universitat de Barcelona, Diagonal 647, E–08028 Barcelona, Spain.

(Dated: April 24, 2007)

We present a theoretical framework for characterizing the geometrical properties of the space of solutions in constraint satisfaction problems, together with practical algorithms for studying this structure on particular instances. We apply our method to the coloring problem, for which we obtain the total number of solutions and analyze in detail the distribution of distances between solutions.

PACS numbers: 05.70.Fh, 89.20.Ff, 75.10.Nr, 02.70.-c

Constraint satisfaction problems (CSPs) offer a unified language describing many complex systems. Originally investigated by computer scientists in relation with algorithmic complexity [1], CSPs have recently attracted much interest within the physics community, following the discovery of their close ties with spin-glass theory [2,3]. They are currently used to tackle systems as diverse as, among others, error-correcting codes [4], rigidity models [5], and regulatory genetic networks [6]. The ubiquity of CSPs stems from their general nature: given a set of discrete variables subject to $M$ constraints, a CSP consists in deciding whether there are assignments of the variables satisfying all the constraints. Of special interest is the class of NP-complete problems [1], for which no algorithm is known that guarantees to decide the satisfiability of a problem instance in a time polynomial in $N$. A well-studied example is the $q$-coloring problem ($q$-COL): given a graph with $N$ nodes and $M$ edges connecting certain pairs of nodes, and given $q$ colors, can we color the nodes so that no two connected nodes share a common color?

Much insight into CSPs is gained by focusing on typical instances drawn from an ensemble with a fixed density of constraints $\alpha = M/N$. As $\alpha$ is varied, a threshold phenomenon is generically observed. Below a critical value $\alpha_c$, instances are typically satisfiable (SAT phase): at least one satisfying assignment (or solution) exists with probability one when $N \to \infty$; above $\alpha_c$ they are typically unsatisfiable (UNSAT phase). Rigorous bounds on $\alpha_c$ have been derived [7]. The running time of algorithms often increases greatly near $\alpha_c$ [8].

CSPs enter the standard framework of statistical physics by associating to each assignment of the $N$ variables $\sigma \equiv \{\sigma_i\}_{i=1}^N$ an energy $E[\sigma]$ defined as the number of constraints violated by $\sigma$. The satisfiability problem reduces to the determination of the ground-state energy $E_0 = \min_{\sigma} E[\sigma]$: if $E_0 > 0$ the instance is UNSAT, if $E_0 = 0$ it is SAT. In recent years, several methods borrowed from statistical physics [2,3,11] have pointed to the existence of a second threshold $\alpha_d < \alpha_c$, associated with clustering of the space $S$ of all solutions. For $\alpha < \alpha_d$ (Easy-SAT phase), $S$ is typically connected: any two solutions are joined by a path of moves involving a finite number of variables. For $\alpha_d < \alpha < \alpha_c$ (Hard-SAT phase), $S$ is typically disconnected: solutions gather into clusters far apart from each other (as illustrated in Fig. 1a), which can only be joined by moves involving a finite fraction of the variables. This scenario, which has been confirmed rigorously in some cases [11], suggests that computational hardness may be caused by the trapping of local algorithms in metastable clusters, which are exponentially more numerous than clusters of solutions.

In this Letter, we introduce methods to analyze in detail the structure of the solution space of CSPs in the Hard-SAT phase. The first aspect we analyze is the entropic structure. A cluster $\lambda$ typically contains an exponential number of solutions, $M_\lambda \approx \exp(N s_\lambda)$, where $N s_\lambda$ is the internal entropy of $\lambda$ (we write $a_N \equiv b_N$ when $\ln a_N/\ln b_N \to 1$ as $N \to \infty$). We introduce the entropic complexity $\Sigma_s(s)$ that counts the number of clusters $N(s) = \exp[\Sigma_s(s)]$ of clusters with internal entropy $N s$, and a method for computing $\Sigma_s(s)$ and the total entropy density $\rho_{\text{tot}}$, yielding the total number of solutions, $|S| \approx \exp(N s_{\text{tot}})$, for individual instances of any CSP. The problem of counting the number of solutions of a CSP is in general #P-complete [12], a class of problems even harder than NP-complete [13]. Estimating $|S|$ is important in applications such as graph reliability [12], and computing partition functions.

A second, related aspect of the structure of $S$ is its geometry. We introduce a method to compute the geometric complexity $\Sigma_d(d)$, which counts the number of clusters at a given distance $Nd$ from a reference assignment (see Fig. 1b), and the related weight enumerator function, of direct interest in coding theory [14]. Finally, we indicate several generalizations of these methods.

Our methods are based on extensions of the “energetic” cavity method (CM) of Ref. [15]. We illustrate them for $q$-COL and show numerical results for $q = 3$, but we emphasize that any CSP can be studied along the same lines. The energy function associated to $q$-COL is that of the antiferromagnetic Potts model, $E[\sigma] = \sum_{i,j} \delta_{\sigma_i,\sigma_j}$, where $\sigma_i \in \{1,\ldots,q\}$ and the sum is over the $M$ graph edges. We study Erdős-Rényi random graphs [16], con-
structured by connecting any pair of nodes with probability $2\alpha/N$. For large $N$ this gives $M = \alpha N$ and a Poisson-distributed connectivity with mean $2\alpha$.

In the unclustered phase (“replica symmetric” phase in the language of spin-glass theory) the zero-temperature energetic CM \[15\] computes the ground state energy $\epsilon \equiv Z_{\sigma_k}(\alpha, \epsilon) = \ln \left[ N_{\text{tot}} \Delta S^{(i)} - \sum_{(i,j)} \Delta S^{(j)} \right]$ where, similarly to the energetic CM \[2\], we need to subtract the link contributions $\Delta S^{(i,j)} = \ln (1 - \sum_{\tau} \bar{\psi}^{(i-j)}_{\tau} \bar{\psi}^{(j-i)}_{\tau})$ from the node contributions $\Delta S^{(i)} = \ln \sum_{\tau} \prod_{k \in i} (1 - \bar{\psi}^{(k-i)}_{\tau})$ to avoid double counting. Above $\alpha_d$, following the 1RSB Ansatz \[22\], we assume the existence of many clusters. We then compute a potential $\phi(x)$ related to the entropic complexity $\Sigma_k(s)$ through

$$e^{N \phi(x)} = \int_{s_{\text{min}}}^{s_{\text{max}}} e^{N \Sigma_k(s)+x s} ds ,$$

where $x$ is a Lagrange multiplier which fixes the internal entropy and $s_{\text{min}}, s_{\text{max}}$ are the points at which $\Sigma_k(s)$ vanishes. Assuming the independence of the quantities $Z_{\sigma_k}^{(i)}$ within each cluster, we introduce probability distributions of the cavity fields $P(x^{(i-j)})(\psi^{(i-j)})$ with respect to the clusters, and generalize the cavity recursion to

$$P(x^{(i-j)})(\psi^{(i-j)}) \propto \int \prod_{k \in i-j} dP(k-i)(\psi^{(k-i)}) Z(\{\psi^{(k-i)}\})^x$$

$$\delta(\psi^{(i-j)} - \bar{\psi}^{(i-j)})(\{\psi^{(k-i)}\}) .$$

After solving Eq. \[3\], the potential is computed as

$$N \phi(x) = \sum_i \ln \int \prod_{j \in i} dP(k-i)(\psi^{(k-i)}) e^{x \Delta S^{(i)}(\psi^{(k-i)})} \Delta S^{(i)}(\psi^{(k-i)})$$

$$- \sum_{(i,j)} \ln \prod_{a \neq (i,j)} dP(a)(\psi^{a}) e^{x \Delta S^{(i,j)}(\psi^{(i-j)})} ,$$

where $\Delta S^{(i)}$, $\Delta S^{(i,j)}$ are given above. A saddle point evaluation of Eq. \[4\] gives $x = -\partial_s \Sigma_k(s)$. Hence, from $\phi(x)$ we obtain $\Sigma_k(s)$ via the Legendre transform $s(x) = \partial_x \phi(x)$, $\Sigma_k(x) = \phi(x) - x s(x)$. We solve numerically Eq. \[3\] on individual graphs by representing the distributions $P(x^{(i-j)})$ with a population of $N \rho$ cavity fields on each oriented link. The resulting message passing algorithm is an entropic generalization of survey propagation \[14\]. Our entropic CM provides greater information at
the price of greater computational difficulty, due to the continuous nature of the cavity fields.

Figure 2 displays some of our results for 3-COL, for three individual graphs with \( N = 5000 \). In particular, the total entropy of solutions is given by \( s_{\text{tot}} = \phi(1) = s_{\text{max}} \) where the last equality holds because, according to our numerical results in Fig. 2, \( \Sigma_s(x) \) vanishes at \( x = x^* < 1 \), with \( s_{\text{max}} = \partial_x \phi(x = x^*) \).

Therefore, for 3-COL the total entropy is dominated by a subexponential number of giant clusters: a randomly chosen solution falls almost surely in one of such rare clusters. We also find that the fraction of frozen variables is finite in the interval \([s_{\text{min}}, s_{\text{max}}]\).

We also implemented a version of Eq. 3 averaged over Erdős-Rényi graphs, by considering a population of links with Poisson connectivity and a population of cavity fields on each link. Figure 3 shows the graph averages obtained in this way for \( s_{\text{tot}} \), the typical internal entropy \( s_{\text{typ}} = \arg \max_s \Sigma_s(s) \), and the typical complexity \( \Sigma_{\text{typ}} = \Sigma_s(s_{\text{typ}}) = \Sigma_s(0) \), as a function of \( \alpha \). The graph-averaged complexity curves \( \Sigma_s(s) \) (not shown) resemble those in Fig. 2 (Graph-to-graph fluctuations for \( N = 5000 \) are significant: for \( \alpha = 2.3 \), the standard deviation of \( \Sigma_{\text{typ}} \) is about 27% of the mean). The negative complexities in Fig. 3 have no direct interpretation on individual graphs, but for the graph-averaged case they are related to rare atypical graphs.

The above formalism can be generalized to yield \( \Sigma_{\text{typ}}(\epsilon, s) \), the complexity associated with metastable clusters of energy \( N\epsilon > 0 \) and entropy \( Ns \), with \( \Sigma_{\text{typ}}(0, s) = \Sigma_s(s) \), by adding a second multiplier \( y \). An equivalent information is contained in the finite temperature complexity \( \Sigma_f(f; \beta) \), where \( f \) is the free energy and \( \beta \) the inverse temperature, based on the identity

\[
\int e^{N[\epsilon_c(\epsilon, s) - \epsilon_s + x s]} \, ds = \int e^{N[\epsilon_f(f; \beta) - x \beta \epsilon_f]} \, df, \quad (5)
\]

with \( f = \epsilon - s/\beta \) and \( y = \beta x \). The energetic CM is recovered for \( \beta \to \infty \) and \( x \to 0 \) with \( y = \beta x \) fixed, which amounts to ignore all entropic effects.

**Counting clusters at a given distance** — We now turn to the geometric structure and show how the CM can be used to investigate inter-cluster distances. We illustrate this by addressing the problem of counting the number of clusters as a function of their distance from a fixed reference configuration \( \varsigma \), which we rephrase as a new CSP, named dCSP, whose thermodynamics reflect the geometry of the solution space of the initial CSP. The valid assignments of dCSP are the solutions \( \sigma \in S \) of the initial CSP; these are configurations of zero energy, and in this sense dCSP concentrates on the zero temperature case of the original problem. But we introduce in dCSP a new energy function which is the Hamming distance \( \varsigma, E_D[\sigma] \equiv \sum_{i=1}^N (1 - \delta_{\varsigma_i, \sigma_i}) \). Therefore the clusters (resp., assignments) of dCSP with energy \( E_D \) are the zero-energy clusters (resp., solutions) at distance \( E_D \) from \( \varsigma \) in the initial CSP.

The optimization problem for dCSP consists in finding the maximal (or the minimal) distance between \( \varsigma \) and a solution of the original problem. By applying the energetic CM to this problem, we obtain a geometric complexity \( \Sigma_d(d) \) giving the number of clusters at distance \( N\delta \) of \( \varsigma \), \( N\delta \approx \exp[\delta \Sigma_d(d)] \). Figure 4 shows results for 3-COL on individual graphs. Two features are worth noticing: i) \( \Sigma(d) \) becomes positive only above a threshold \( d_{\text{min}} \), reflecting the fact that clusters are well separated; ii) a plateau appears between \( d_1 \) and \( d_2 \), reflecting the finite diameter of clusters. We have verified

\[
\int e^{N[\epsilon_c(\epsilon, s) - \epsilon_s + x s]} \, ds = \int e^{N[\epsilon_f(f; \beta) - x \beta \epsilon_f]} \, df, \quad (5)
\]
that the size of this plateau coincides with the typical diameter computed within the entropic CM \[25\].

Generalizations – The above method can be extended to count the number of solutions at distance \(N d\) from \(\varsigma\), known as the weight enumerator function \(A_N(d)\) in coding theory \[14\]. This can be deduced from the complexity \(\Sigma_{d,s}(d,s)\) which gives the number of clusters with internal entropy \(N\varsigma\) at distance \(N d\) from the reference configuration \(\varsigma\). Such a complexity can be obtained by studying the dCSP with a finite value of a new inverse temperature \(\beta_d\), which is conjugate to the energy \(E_d\) (keeping the original temperature \(\beta^{-1}\) to zero) \[25\]. Once \(\Sigma_{d,s}(d,s)\) has been found, one obtains the leading behaviour of the weight enumerator as \(A_N(d) \approx \exp[N \max_s(\Sigma_{d,s}(d,s) + s)]\). In the same spirit, our analysis can be extended to metastable configurations: in order to compute the complexity \(\Sigma_{c,d,s}(e,d,s)\) counting clusters with energy \(N\varsigma\), entropy \(N\varsigma\), at distance \(N d\) from \(\varsigma\), one needs to introduce three Lagrange multipliers \(x, y, z\). All the previous complexities are particular limits of this more general framework \[25\].

Conclusions – We have presented methods to analyze the entropic and geometric structure of the clustered phase in \(q\)-COL, which give access to quantities such as internal cluster entropies not accessible to previous methods. Our results for 3-COL show the existence of giant, atypical clusters which contain the majority of solutions. Generalization to other CSPs such as \(k\)-SAT, where a similar picture may hold, is straightforward.

Notice that the present results were obtained within a 1RSB ansatz, and the stability of our solution should thus be checked (extending the method of Ref. 20) to assess whether the solution is exact or only an approximation to a more complicated one involving higher order RSB.

The new information extracted with our entropic CM could be exploited to design new algorithms for finding solutions to individual instances, improving on present survey propagation algorithms which only use energetic information \[11\]. We also envision applications to inference problems such as Bayesian belief networks \[28\].

We thank D. Battaglia and R. Zecchina for discussions, and A. Pagnani for sending the SP code used in Refs. 17, 18. This work was supported in part by the European Community’s Human Potential Programme under contracts HPRN-CT-2002-00319 (STIPCO) and by the Community’s EVERGROW Integrated Project.

Note added: the recent paper \[29\] addresses similar questions.

\[\text{FIG. 4: Geometric complexity for the same individual graphs as in Fig. 2. Within the error bars, } \Sigma_d(d_1) = \Sigma_d(d_2) = \Sigma_s(s_{typ}). \text{ The horizontal lines are a guide to the eye.}\]
with the derivative of a polynomial fit of $\phi(x)$.


[26] When $\beta \to \infty$ at fixed $x$, the cavity field $\psi(\beta)$ used in computing $\Sigma_f(f;\beta)$ in Eq. (4) yields either the purely entropic field $\psi = \lim_{\beta \to \infty} \psi(\beta)$ or the energetic field $h = \lim_{\beta \to \infty} \{\ln[\psi(\beta)]/\beta\}$ (the value of $x$ is irrelevant in the latter case). The supplement of entropic information conveyed by $\psi$ defines the “evanescent field” $h' \equiv \lim_{\beta \to \infty} \{\ln[\psi(\beta)] - \beta h\}$.

[27] If $\zeta \in S$, the minimal distance is zero, but by focusing on the metastable (glassy) states, the cavity method allows to ignore this trivial (crystalline) ground-state.


[29] A. Bandyopadhyay and D. Gamarnik, math.PR/0510471.