Systematic Generation of Very Hard Cases for Graph 3-Colorability

Romulus Dan VLASIE
I3S Laboratory
University of Nice
Sophia Antipolis, FRANCE, 06903

Abstract

We present a simple generation procedure which turns out to be an effective source of very hard cases for graph 3-colorability. The graphs distributed according to this generation procedure are much denser in very hard cases than previously reported for the same problem size. The coloring cost for these instances is also orders of magnitude bigger. This ability is issued from the fact that the procedure favors -inside the class of graphs with given connectivity and free of 4-cliques- the generation of graphs with relatively few paths of length three (that we call 3-paths). There is a critical value of the ratio between the number of 3-paths and the number of edges, independent of the number of nodes, which separates the graphs having the same connectivity in two regions: one contains almost all graphs free of 4-cliques, while the other contains almost no such graphs. The generated very hard cases are near this phase transition, and have a regular structure, witnessed by the low variance in node degrees, as opposite to the random graphs. This regularity in the graph structure seems to confuse the coloring algorithm by inducing an uniform search space, with no clue for the search.

1 Introduction

Providing tools for the systematic generation of test cases for benchmarking A.I. programs does not lack of interest. Such tools would provide a common reference for people involved in developing and testing new search algorithms, allowing evaluation and comparison of different heuristics.

An evident requirement for the test data is their *hardness*, in the sense that the data must be difficult for the tested algorithms. Beside the practical importance for the experimental evaluation of algorithms, coping with the systematic generation of test cases is a good occasion to take a closer look at the structure of hard instances. Decrypting the patterns of hard cases allows their automatic generation, and may give also hints for the design of more appropriate algorithms.

Since most A.I. applications reduce to the solving of NP-complete problems, it seems that the hardness issue is guaranteed (assuming $P \neq NP$). For instance, it was shown ([7]) that for the Tiling problem, the simple, uniform selection of instances leads to superpolynomial on average computations (unless there are polynomial on the average algorithms for every NP-complete problem and every simple distribution of its instances). On the other hand, for many NP-complete problems, such as K-sat, graph K-colorability, etc., finding instances requiring a superpolynomial time is not straightforward. For example, in [13] it is shown that random K-colorable graphs, with K small, can be colored with high probability in polynomial time. In fact, there is an algorithm, the Brélaz algorithm, which runs in $O(\mu^2)$ time on almost all K-colorable graphs with μ vertices ([1], [13]).

Searching for hard cases

A source of hard instances for many NP-complete problems was identified in [2], by conjecturing that the hard instances are around a critical value of an *order parameter* which separates by an abrupt *phase transition* the region where almost all instances are solvable from the region where almost all instances have no solution. Much empirical and theoretical evidence in this sense has been given for graph coloring ([2],[14]), K-satisfiability ([10],[3]), random CSPs ([11]). The good news with this approach is that by simply reading a "magic number", it is possible to predict whether a given case belongs to the easy or the hard region of the space of instances. For example, it was found experimentally that for graph 3-colorability there is a peak in the median search cost for random graphs having the critical connectivity (average degree) $\gamma \approx 4.6$ ([5]), or around $\gamma \approx 5.4$ for random reduced graphs ([2]).

This does not exclude the existence of hard cases elsewhere, outside the critical region. In fact, in [6] the authors report that harder cases of graph 3-colorability occur in a supposed easy region, for random graphs having different connectivities, below the critical value. For convenience, we call this region the range of the *second phase transition*; in this range of connectivities the scaling of the mean cost changes from being polynomial to being exponential, as argued in [6]. Similar results about a second hard region are reported for satisfiability problems ([4]) and random CSPs ([12]). All these cases are significantly harder than the previously observed instances from the critical region, and it has been proposed to use them for benchmarking search algorithms. However, there are several drawbacks with the cases from this second hard region. Firstly, these cases are very rare, lost inside the mass of trivial instances. For example, in [6] the authors explored up to one million samples of 50-node graphs with the connectivity near $\gamma = 3$, in order to find about 50 cases with the cost bigger than 5000 (the cost is expressed as being the number of search states expanded by the Brélaz algorithm). Secondly, these hard instances seem to lack any obvious syntactical criterion which eventually could allow their automatic generation. And thirdly, by simply changing the search heuristics or the problem presentation, these instances often become very easy ([15],[4]), suggesting that there is a question of "poor algorithms" rather than of "intrinsically hard" problems[1].

Paper contribution

Our belief is that the $P \neq NP$ assumption is strong enough to allow the existence of regions in the space of instances containing with high probability instances much harder that the ones observed so far. In this paper we provide abundant experimental evidence in this sense for graph 3-colorability, by proposing a certain generation procedure which is very effective in producing extremely hard cases. These very hard graphs to color lie in a range of connectivities comparable to the range locating the second phase transition, but inside each region of given connectivity the very hard cases are concentrated following a second structural parameter.

Namely, we associate to each graph G with μ nodes and m edges[2] the number of its *3-paths*, as a more subtle measure of the graph's topology. To precise, we call 3-path any path in the graph having the length equal to three (denoted by an alternate succession of vertices and edges $x_1e_1x_2e_2x_3e_3x_4$, $x_1 \neq x_4$)[3]. We denote the number of 3-paths in a graph G by $\pi(G)$. With these precisions, we found that for the graphs in the range of the second phase

[1] This is a suggestion of Patrick Prosser made on the newsgroup *comp.constraints*.

[2] All the graphs with μ nodes and m edges have the same connectivity $\gamma = \frac{2m}{\mu}$.

[3] We will not distinguish a 3-path from its reverse path, such that both $x_1e_1x_2e_2x_3e_3x_4$ and $x_4e_3x_3e_2x_2e_1x_1$ will denote the same 3-path, being counted as one.

transition, generated according to our procedure, the very hard cases tend to concentrate near a critical value π_{crit} of the number of 3-paths (the number of nodes and the graph connectivity are kept fixed). This critical value marks the abrupt decay (phase transition) in the number of graphs free of 4-cliques and having a prespecified number of 3-paths π: the graphs free of 4-cliques with $\pi > \pi_{crit}$ are exponentially many, while those with $\pi < \pi_{crit}$ are exponentially few. Furthermore, by varying the number of nodes μ, but not the graph connectivity, we found that the value of the ratio between the critical number of 3-paths and the number of edges π_{crit}/m remains the same, showing that the ratio π/m is a new order parameter for graph 3-coloring.

In the next section we describe in detail the procedure that we used to generate graph instances. The experimental results obtained by using the Brélaz algorithm are discussed in Section 3. The paper ends with references at related work and with a final section of conclusions.

2 Generation of very hard graphs

The graph 3-colorability problem consists in answering *yes/no* if the nodes of a given graph can/cannot be colored with three colors, such that no two adjacent nodes have the same color. Our hypothesis is that different edge placements discriminate between easy and hard instances having the same number of nodes and edges. If this hypothesis is true, it could be possible to generate hard cases systematically by reproducing the corresponding edge configurations.

In order to be aware of what these edge configurations could be, we propose a *generating process* which proceeds by small random changes in the edge configuration, starting from a random graph. The occurrence of 4-cliques is avoided, as being a trivial source of problems with no solution.

A generating process of parameters μ and m denote a finite series of graphs $G_0, \ldots, G_n$, such that:

- all $G_i, i = 0..n$ have μ vertices and m edges;
- all $G_i, i = 0..n$ are free of 4-cliques;
- the difference between two consecutive graphs G_i, G_{i+1} is given by the placement of only one edge;
- the number of 3-paths is strictly decreasing : $\pi(G_i) > \pi(G_{i+1}), i = 0..n-1$;
- (Endingpoint Condition) there is no possibility to extend the graph series $G_0, \ldots, G_n$ with a graph G_{n+1} by verifying all the above conditions.

We call G_0 *the starting point* and G_n *the ending point* of the generating process, while n is *the length* of the process.

We implement such a generating process by running a simple procedure (see Figure 1) which starts by choosing G_0 as random with μ vertices, m edges, and free of 4-cliques. Then the procedure successively generates the graph G_{i+1} from G_i by randomly removing one edge and randomly drawing another one, such that no 4-clique is formed and the number of 3-paths is strictly decreasing. The procedure stops with G_n as ending point, when there is no possibility to generate a new graph G_{n+1} from G_n.

There is a simple, incremental formula to count the number of 3-paths in a graph: if one edge xy is added/removed to/from the graph G, then the number of 3-paths $\pi(G)$ is

```
procedure GENCOL /* GENerating 3-COLoring test cases */
repeat
   take G_0 as random with μ vertices and m edges;
until G_0 has no 4-cliques or more than 1000 graphs was tried;
if more than 1000 graphs were tried then
   raise exception("4-cliques are very probable");
put i := 0;
repeat
   put endingpoint :=true;
   for each edge e ∈ Edges(G_i) do /* random order of edges */
      for each pair (x, y), x ≠ y, of vertices of G_i do /* random order */
         if edge(x, y) ∈ Edges(G_i) then
            continue with another pair;
         /* construct a new graph G */
         put Vertices(G) := Vertices(G_i)
         put Edges(G) := Edges(G_i) \ {e} ∪ {edge(x, y)}
         if G has no 4-cliques and π(G) < π(G_i) then
            put i := i + 1;
            put G_i := G;
            put endingpoint :=false;
            break both the for-loops;
      endfor
   endfor
until endingpoint;
endprocedure;
```

Figure 1: *The procedure implementing a generating process.*

increased/decreased with the following quantity:

$$\Delta\pi(G) = \sum_{v \in N(x)} d(v) + \sum_{v \in N(y)} d(v) + d(x)d(y) - d(x) - d(y) - 3\,|N(x) \cap N(y)|, \qquad (1)$$

where x, y are the ends of the edge, $d(v)$ means the degree of the vertex v and $N(x), N(y)$ are the sets of immediate neighbors of the vertices x and y respectively, considered after/before the edge xy is added/removed.

Claim 1 *The construction of the graphs $G_0, \ldots, G_n$ of a generating process is done in polynomial time.*

Proof: Formula 1 shows that the variation in the number of 3-paths of a graph is polynomial (bounded by $O(\mu^2)$) and then the total number of 3-paths in a graph is polynomial following the graph size (bounded by $O(\mu^2)$). Since for each generated graph the number of 3-paths is strictly decreasing, it follows that the generating process has a polynomial length. A graph is generated after the execution of the two nested **for** loops from Figure 1. This execution takes at most m times $\mu\,(\mu-1)/2 - m$ steps, and that when the endingpoint is detected. Each step consists in a 4-clique test and a computation of the number of 3-paths; the 4-clique test is done in linear time by exploring the immediate neighbours of nodes, and the computation of the number of 3-paths is done incrementally by applying Formula 1, thus not affecting the polynomial scaling of the entire generating process. With these considerations the proof is achieved.
◇

3 Experimental results

We provided sample sets for each connectivity in the range $\gamma = 2.4 \ldots 6$, given in increments of 0.4. For each connectivity we collected in a sample set the graphs obtained by running 100 independent generating processes. The data were generated for 50-node graphs and for 80-node graphs.

Measuring the hardness

The difficulty of the constructed graphs is registered by using the Brélaz algorithm ([1]). A number of papers ([13], [9], [2], [15]) shows that this algorithm still remains largely unrivaled for graph coloring. In particular, the results about the existence of the first phase transition ([2]) and the second phase transition ([6]) are established using the Brélaz algorithm. This algorithm is a complete backtracking procedure, which use heuristics: at each step an uncolored node with fewest remaining colors is chosen; ties are broken by selecting the node with maximal degree in the subgraph of uncolored nodes; remaining choices are made randomly. We optimized the search by a preliminary step which decompose the graph in connect components, then the Brélaz algorithm is applied to each component; this saves unnecessary backtracking when there is an uncolorable component. As a direct measure of difficulty we use the number of search states expanded by the Brélaz algorithm, where the search finds the first solution or proves that none exists. A new state is expanded each time a node changes its color during the search. The preprocessing step of decomposing in connect components is not counted, since it is achived in linear time and is systematically done for each graph. In order to avoid costly computations, we stopped the searches exceeding three millions Brélaz steps, by considering this bound as illustrative for the search difficulty of 50-node and 80-node graphs.

Generation effectiveness

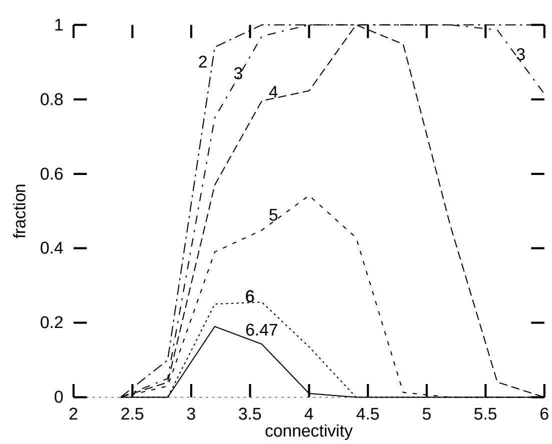

Figure 2: *The experimental probability that a generating process for 50-node graphs constructs an instance demanding a cost greater than a given value. Near each curve we marked the* $\log_{10}$ *of the exceeded cost. For each connectivity we ran 100 independent processes.*

The first issue that we addressed on the obtained results is about the effectiveness of the generation procedure GENCOL in producing hard graphs. In this sense we measured the fraction of generating processes which construct at least one graph with the search cost exceeding a given value. The results are showed in Figure 2, where each curve represents the experimental probability that a generating process produces a 50-node graph with the $\log_{10}$ of the cost exceeding the label of the curve. For instance, we can see that for the connectivities in the range $3.6 \ldots 5.6$ the probability to generate a 50-node graph with the cost exceeding 1000 is practically one. Also, the chances to generate a 50-node graph with $\gamma = 3.2$ and whose cost is exceeding one million Brélaz steps are about 25%, while there are about 19% chances that the cost exceeds even three millions.

Mean cost

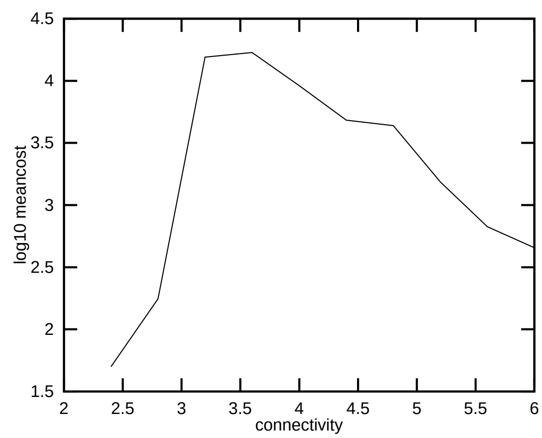

Figure 3: *Mean cost vs. connectivity for 50-node graphs, under the distribution implied by the generation procedure.*

Another experimental issue concerns the mean cost under the particular distribution imposed by the generation procedure. Figure 3 plots the mean cost for each sample set of given connectivity, when the number of nodes is 50. Each sample set was obtained by running 100 independent generating processes. The results are underestimated by the fact that there are searches stopped when the cost is exceeding three millions steps. However, we clearly see that the mean cost peaks for an intermediary connectivity between 3.2 ... 3.6. The peak becomes sharper when 80-node graphs was considered, but its location does not change.

Comparying with the uniform generation

Let us take a closer look at the density of hard instances inside the sample sets. For that purpose we plotted in Figure 4 the experimental cost distribution as percentile curves. The cost C_p corresponding to the curve labeled $p\%$ means that $p\%$ of the costs in the sample set exceed C_p. For the sake of comparison, we reproduce in Figure 5 the experimental cost distribution obtained by Hogg & Williams ([6]) for uniform random 50-node graphs, by using the same Brélaz algorithm. First of all, we remark that our plots (Figure 4) are qualitatively conform with the results reported for uniformly distributed samples (Figure 5): the median cost (50%-th percentile) peaks for γ near 5 (i.e. the first phase transition), while successive higher percentiles peak at different lower connectivities, defining the second phase transition.

But there is a clear quantitative difference in both density and difficulty: for example, the top 0.001 of our generated graphs with γ near 3.5 have the cost exceeding three millions

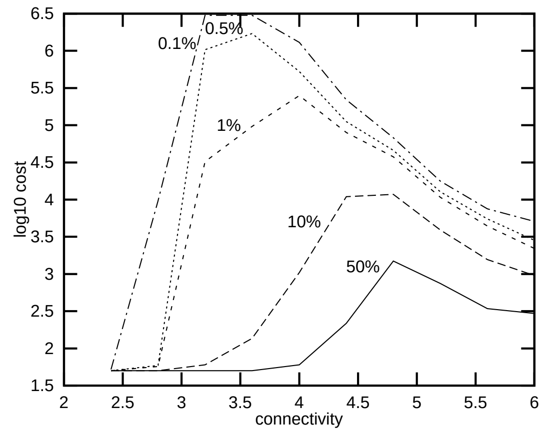

Figure 4: *Cost percentiles vs. connectivity for 50-node graphs obtained by running 100 independent processes at each connectivity given in increments of 0.4.*

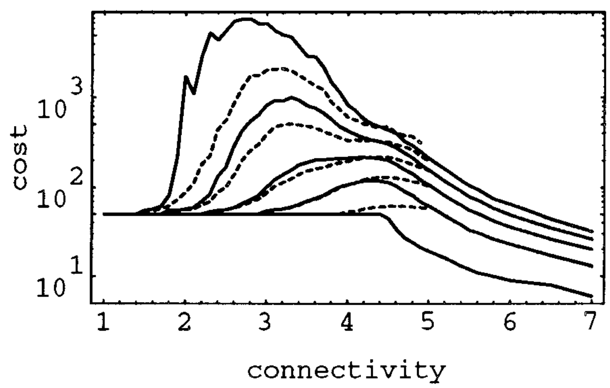

Figure 5: Cost percentiles vs. connectivity for 50-node graphs obtained by uniform random generation.

search steps, while only the top 0.00005 of the graphs uniformly distributed exceed the cost of 5000 steps, for the same number of nodes and the same connectivity.

Locating the very hard graphs

From Figure 3 and Figure 4 we deduce that the cost variance is very large, especially inside the sample sets having the connectivities near the location of the peak in the mean cost. The mean cost is dominated by the few, very hard cases. The question is to see if it is possible to locate the very hard cases inside the sample sets of same connectivity. The natural structural parameter to take into account is the number of 3-paths, as implied by the generation procedure.

Figure 6 contains scatter plots of costs along the different values of the ratio between the number of 3-paths and the number of edges π/m. We considered the sample sets having the connectivities near the location of the peak in the mean cost. We marked the average values of the ratio π/m over the starting points and the ending points respectively, for the 100 generating processes which make the sample set. The horizontal axis is inversed, according to the evolution sense of the generating processes. For each figure, the solid curve shows

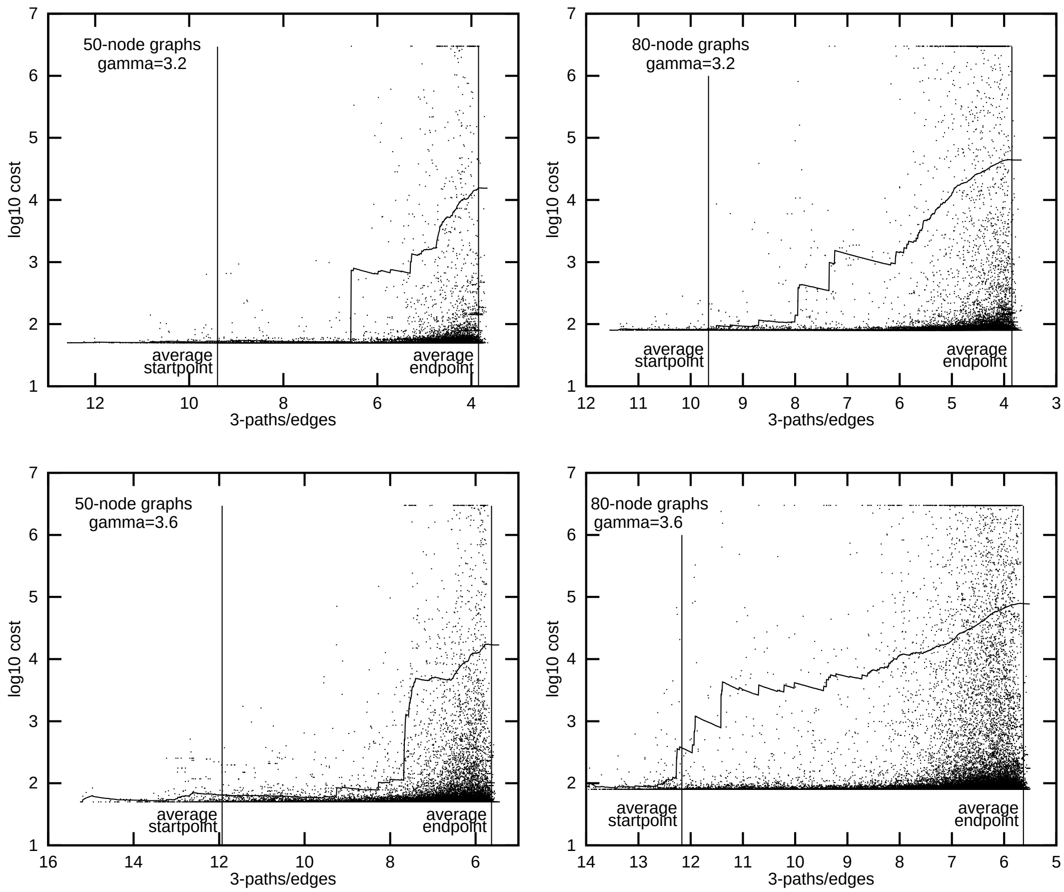

Figure 6: *Structure of the average cost vs. number of 3-paths/edges. We remark the concentration of very hard cases near the value characterizing, on average, the endingpoints. This value separates the region containing many graphs free of 4-cliques from the region containing few such graphs.*

the structure of the mean cost, by indicating the average cost over the graphs generated so far. We clearly see that the average cost is highly dominated by the samples having the ratio π/m near the value characterizing on average the endingpoints. The location of very hard cases inside the sample set does not change when the number of nodes is increased, as showed by the plots from Figure 6

A phase transition

We also see that there are very few graphs whose number of 3-paths is below the value averaged over the endingpoints. This suggests that the generated very hard cases are concentrated near the point π_{crit} marking the phase transition in the number of graphs free of 4-cliques and having a given number of 3-paths π: the graphs with $\pi > \pi_{crit}$ are exponentially many, while those with $\pi < \pi_{crit}$ are exponentially few. The experimentally obtained locations are $\pi_{crit}/m \approx 3.85$ for $\gamma = 3.2$, and $\pi_{crit}/m \approx 5.63$ for $\gamma = 3.6$ The analysis of sample sets with different connectivities showed that the above phenomenon about the concentration of hard cases is less important as the connectivities increase (i.e. the mean cost decreases, as indicated in Figure 3).

Accompanying phenomena

We conclude this section by reporting some accompanying phenomena when generating graphs using the procedure GENCOL. First of all, we remark the dramatic variation in global behavior when small, local changes are operated in the instance structure. Often, by simply redrawing one edge, the coloring cost fluctuates with orders of magnitude. Another important observation is that as the number of 3-paths decreases, so too does the variance of node degrees, the very hard instances having a lower variance of node degrees. A question of further investigations is to see if the variance of the degrees is an order parameter per se, independent of the number of 3-paths. Finally, we note that we found no corelation between problem solvability and problem difficulty inside the sample sets having the connectivities in the range of the second phase transition.

4 Related work

We seen that the strictly decrease in number of 3-paths is accompanied by a decrease in the variance of node degrees. The very hard instances, characterized by a small number of 3-paths, have also a small variance of node degrees. In [2] the authors also considered the possibility that the variance of node degrees could be an additional parameter to locate the hard instances of graph colorability, concluding that graphs with lower variance of the degrees are somewhat harder to color. But since they analyzed only the random reduced graphs located inside the first phase transition (i.e. the average degree $\gamma \approx 5$), they concluded that the variance of the degrees is only a weak additional order parameter. The variance of node degrees for graphs taken inside the second phase transition was analyzed in [15], but since the samples was uniformly distributed, no relation between the search difficulty and the variance of the degrees was found.

A systematic method proposing additional structural parameters for graph coloring is discussed in [5]. Based on the pruning effect of particular edge configuration, this method indicates that the number of triangles, squares, etc. in the graph are order parameters also. In the same paper the coloring difficulty for graphs selected under various distributions is also analyzed. In particular, the "padded" graphs, having the structure close to random regular graphs, were found as being much harder to color than uniform random ones.

5 Conclusion and further work

In this paper we presented a generation procedure which turns out to be an effective source of very hard cases for 3-colorability problem. The frequency of very hard cases is sensibly higher and the colorability cost is several orders of magnitude bigger than previously recorded ([6]). This ability is issued from the fact that the procedure favors -inside the class of graphs with the same connectivity, free of 4-cliques- the generation of instances with relatively few 3-paths.

There is a value π_{crit} such that the graphs free of 4-cliques and having the number of 3-paths π greater than π_{crit} are exponentially many, while those having the number of 3-paths smaller than π_{crit} are exponentially few. We experimentally found that, in the range of the second phase transition, the average cost over the generated samples is heavily increased by the graphs having the number of 3-paths π near π_{crit}.

Graphs with relatively few 3-paths have a regular structure (witnessed by the lower variance of node degrees), as opposite to random graphs, which have many 3-paths. A legitimate question still remains: why the more regular instances, with more uniform structure, tends to be much harder than the random instances? The answer could be that these regular instances induce uniformity in the structure of the search space, offering no clue to the search algorithm, confused by the equally promising choices.

Altough we experimented with only the Brélaz algorithm, we conjecture that the same behavior could be observed for *any* other comparable graph coloring algorithm. We already started the development of a formal framework backing these affirmations. The main issue is based on a theorem due to Li and Vitányi saying that the average case complexity under the *universal distribution* equals the worst-case complexity ([8]). This immediately implies that under the universal distribution, *any* algorithm solving an NP-complete problem will run on average in a superpolynomial time (unless $P = NP$). Preliminary studies show that the distribution of graphs generated according to procedure GENCOL approaches a universal-like distribution, by favoring the occurrence of graphs with regular structure.

References

[1] D. Brelaz. *New methods to color the vertices of a graph.* *Comm. ACM*, (22):251–256, 1979.

[2] P. Cheeseman, B. Kanefsky, and W.M. Taylor. *Where the Really Hard Problems Are.* In Morgan Kaufmann, editor, *Proceedings of IJCAI-91*, pages 331–337, 1991.

[3] J.M. Crawford and L.D. Auton. *Experimental results on the crossover point in satisfiability problems.* In *Proceedings of AAAI-93*, 1993.

[4] I.P. Gent and T. Walsh. *Easy problems are sometimes hard.* *Artif. Intell.*, (70):335–345, 1994.

[5] T. Hogg. *Applications of Statistical Mechanics to Combinatorial Search Problems.* Research Paper, October 1994. Xerox Palo Alto Research Center.

[6] T. Hogg and C.P. Williams. *The hardest constraint problems: a double phase transition.* *Artif. Intell.*, (69):359–377, 1994.

[7] L. A. Levin. *Average Case Complete Problems.* *SIAM J. Comput.*, (15):285–286, 1986.

[8] M. Li and P. M.B. Vitányi. *Average case complexity under the universal distribution equals worst-case complexity.* *Inf. Proc. Letters*, (3):145–149, 1992.

[9] S. Minton and al. *Minimizing conflicts: a heuristic repair method for constraint satisfaction and scheduling problems.* *Artificial Intelligence*, (58):161–205, 1992.

[10] D. Mitchell, B. Selman, and H.J. Levesque. *Hard and easy distributions of SAT problems.* In *Proceedings of AAAI-92*, 1992.

[11] P. Prosser. *Binary Constraint Satisfaction Problems: Some are Harder than Others.* In A.G. Cohn, editor, *Proc. of ECAI-94*, pages 95–99, 1994.

[12] B.M. Smith. *In Search of Exceptionally Difficult Constraint Satisfaction Problems.* In M. Meyer, editor, *Proc. of Constraint Processing Workshop at ECAI'94*, pages 79–86, August 1994.

[13] J. S. Turner. *Almost All k-Colorable Graphs Are Easy to Color.* *Journal of Algorithms*, (9):63–82, 1988.

[14] C. P. Williams and T. Hogg. *Exploiting the deep structure of constraint problems.* *Artificial Intelligence*, (70):73–117, 1994.

[15] C.P. Williams and T. Hogg. *Hard Problems Are Almost Everywhere.* Research Paper, 1993. Xerox Palo Alto Research Center.