1E1-R-1-3

グラフ彩色インスタンスの組織的生成のための極小非可解構造の導出

Finding Minimal Unsolvable Structures for Constructive Generation of Graph 3-Colorability Instances

水野 一徳[*1] 西原 清一[*2] 佐々木 整[*1]
Kazunori Mizuno Seiichi Nishihara Hitoshi Sasaki

[*1]拓殖大学 工学部 情報工学科
Department of Computer Science, Takushoku University

[*2]筑波大学大学院 コンピュータサイエンス専攻
Department of Computer Science, University of Tsukuba

We have developed the method for systematically generating very hard graph 3-colorability (3COL) instances to clarify phase transition mechanism of combinatorial search problems. In our method, 3COL instances can be generated by repeatedly embedding minimal unsolvable graphs (MUGs). In this paper, we find larger-scale MUGs, which are necessary for our generation method, using genetic algorithms (GAs). We also experimentally demonstrate that the computational cost to solve our 3COL instances generated by using MUGs found by GAs is of an exponential order of the number of vertices and our instances are extraordinarily harder than randomly generated instances.

1. はじめに

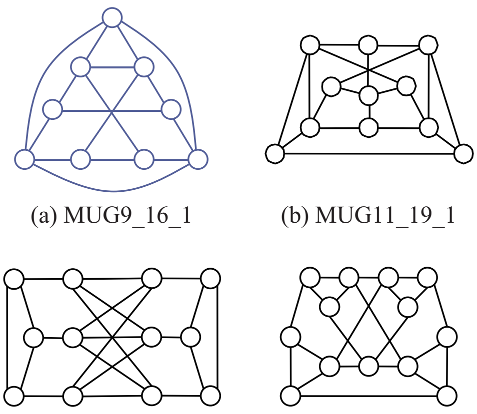
グラフ彩色問題（Graph Colorability: COL）は，制約充足問題（CSP）の代表的な例題として解探索アルゴリズムの開発や現実問題への応用など幅広く研究されている．CSP と同様に相転移（Phase Transition: PT）現象 [Hogg 96] が発生することが観察されており，ヒューリスティクス [Brélaz 79] の研究題材としても興味深い．グラフの頂点に塗る色数を 3 とした COL（3COL）は NP 完全であり，多くの研究報告が存在する．解くのに手間のかかる 3COL の特徴として，Vlasie の 3-path[Vlasie 95]，Mammen らの極小非可解部分問題 [Mammen 97]，Culberson らの frozen development[Culberson 01] など多くの有効なパラメータが提案されている．しかし，これらの研究は生成検査に基づく方法，つまりランダムにインスタンスを生成して難しいものを選びとるアプローチであるため，安定的にそのような難しいインスタンスを得ることが難しい．

これに対して，文献 [Mizuno 08, Nagasawa 09] では，それらの難しいインスタンスが持つ構造条件を備えた構造（n4c-free MUG）を提案し，その構造を用いて難しい 3COL インスタンスを組織的に生成する方法を提案している．文献 [Mizuno 08, Nagasawa 09] の方法でインスタンスを生成するには，n4c-free MUG 構造が重要であり，それらが多様なほどより複雑なインスタンスが生成できる．しかし，文献 [Mizuno 08, Nagasawa 09] では，グラフ集合に対して試行錯誤によってこのような構造を導出しており，より大きなサイズの n4c-free MUG を導出することは困難となってしまう．

そこで，本研究では，GA を用いて n4c-free MUG 構造の導出を試みる．これにより，従来の試行錯誤的な導出よりも効率的に，より多様な，かつより大きなサイズの n4c-free MUG 構造が発見できると考えられる．また，本方法によって導出された n4c-free MUG を文献 [Mizuno 08] で述べられている生成手続きに適用して 3COL インスタンスを生成し，それらが従来より報告されている難しいインスタンスが持つ諸条件を維持していること，およびそれらを解く計算量が頂点数の指数オーダとなることを実験的に示す．

本研究は，PT 領域に属するような難しいインスタンスの持つ構造条件を，そのようなインスタンスを意図的に生成することを通して明らかにしていこうとする試みである．また，今後の組合せ探索，組合せ最適化のアルゴリズム開発において，アルゴリズムの性能評価を行なうためのベンチマークとしての意義も大きいと考えられる．

連絡先: 水野一徳，拓殖大学工学部情報工学科
〒 193-0985 東京都八王子市館町 815-1
tel./fax: 042-665-4789
e-mail: mizuno@cs.takushoku-u.ac.jp



(a) MUG9_16_1 (b) MUG11_19_1 (C) MUG12_21_5 (d) MUG13_22_2

図 1: 極小非可解グラフ（n4c-free MUG）構造の例

2. 研究分野の概要

2.1 グラフ彩色問題と相転移

COL とは，無向グラフの隣接する頂点が同じ色にならないように，すべての頂点に色を塗り分ける問題である．本研究では頂点に塗る色の個数を 3 色とした COL（3COL）を対象とし，色を塗るグラフ G を，$G = (V, E)$ と定義する．ただし，V は頂点集合，E は辺集合とし，$n = |V|$，$m = |E|$ とする．

COL は CSP の代表的な例題である．ここで，ランダムに生成し た CSP 集合について計算量を調べてみると，各イン

スタンスを解決するのに要する平均的な手間は，問題サイズを固定したとき，制約の数が増えるにしたがって易-難-易 (easy-hard-easy) のパターンを示すことが知られている．このとき，計算量の中央値が最も大きくなるのは，可解／非可解の問題がともに 50%の確率で生成される極端に狭い領域付近である．この現象は第 1PT と呼ばれている [Hogg 96]．また，第 1PT よりも制約密度がやや低い領域に おいて,「例外的に非常に計算で手間を要する問題 (Exceptionally Hard Instance: EHI)」が発生する傾向があることが知られている（第 2PT）[Hogg 94].

このようなの現象の存在については明らかになりつつあるものの，その発生原因やメカニズムについては完全には明らかになっていない．Mammen らは， 相転移領域において，インスタンスに含まれる極小非可解部分問題 (Minimal Unsolvable Subproblem : MUS) のサイズが，そのインスタンスサイズに対して大きくなることによって引き起されていると説明している [Mammen 97] ．さらに，ランダムに生成した 3COL の実験では，第 2PT における EHI はそのほとんどが非可解な問題である [Hogg 94]．また，Vlasie は頂点の次数の分散が小さい準正則なグラフほど困難になる傾向があると報告している [Vlasie 95].

2.2 極小非可解構造に基づく組織的生成

2.1 節で述べたインスタンスの難しさの特徴は有効なパラメータであるが，これらはランダムな生成検査に基づくものであり，難しいインスタンスを 安定的に得ることは難しい． これに対して，文献 [Mizuno 08] では，これらの難しい条件を備えたグラフ（n4c-free MUG）集合，および極小非可解性や準正則性を維持したまま n4c-free MUG を繰り返し埋め込むことによって 3COL インスタンスを生成する方法を提案している．また，文献 [Nagasawa 09] では，これをさらに拡張し，新たに 2,000 個以上の n4c-free MUG 構造を提案している．

図 1 は，文献 [Mizuno 08, Nagasawa 09] によって提案された n4c-free MUG の一例を表わしている．この n4c-free MUG 構造は，下記のような特徴がある．

性質 1 極小非可解構造（非可解であるが，任意の辺を 1 本削除すると可解になる構造）である（3COL においては，4-critical graph）.

性質 2 準正則（頂点の次数が 3 か 4 のみで，その分散が小さい）である.

性質 3 n4c 構造（4 頂点クリークから辺を 1 本削除した構造）を部分グラフとして含まない.

ここで，n4c（near 4-clique）とは，4 頂点からなるクリーク（K_4）から辺を 1 本削除した構造であり，frozen pair[Culberson 01] という同じ色にならなければならないという頂点組をもつ特徴的な制約構造を持つものである．図 2 は，文献 [Mizuno 08] で提案されている生成手続きを表わしている．初期グラフ（図 2 の (1)）および埋め込み対象のグラフを先の n4c-free MUG 集合の中から選ぶ（(図 2 の (2)）ことによって，この生成手続きは，極小非可解性，準正則性を維持したまま，かつ n4c を部分構造として含まないという難しいインスタンスの上記 3 つの条件を備えた 3COL インスタンスのサイズを任意に大きくすることができる．実際，文献 [Mizuno 08, Nagasawa 09] では，この方法によって生成されたインスタンスを解く計算量は，グラフの頂点数の指数オーダになることが実験的に示されている．

procedure graph-generator(k)
begin
  input an initial graph G_{init}; (1)
  $G := G_{init}$;
  for $w := 1$ **to** k **do**
    choose randomly an edge$(i,j) \in E(G)$,
      where deg$(i) \leq 3$;
    choose randomly MUG$_{n,t}$; (2)
    embed_MUG$_{nt}(i,j)$;
  end for;
end.
procedure embed_MUG$_{nt}(i,j)$
begin
  choose randomly an edge$(x,y) \in E(MUG_{n,t})$,
    where deg$(x) \leq 3$;
  remove the edge (i,j)
  remove the edge (x,y);
  add an edge (j,y);
  merge x with i;
end.

図 2: n4c-free MUG の埋め込みアルゴリズム [Mizuno 08]

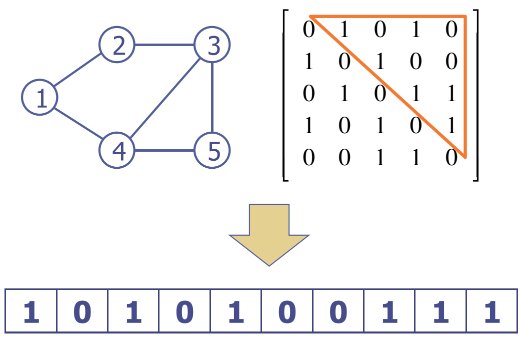


図 3: 隣接行列を用いたグラフ G のコード化

3. GA を用いた極小非可解構造の導出

3.1 概要

文献 [Mizuno 08, Nagasawa 09] の方法でインスタンスを生成するには，図 1 のような構造が重要であり，それらが多様なほどより複雑なインスタンスが生成できる，文献 [Mizuno 08] では，試行錯誤によってこのような構造を導出しており，より大きなサイズの n4c-free MUG を導出することは困難となってしまう．そこで，本研究では，GA を用いることにより，より多様な n4c-free MUG 構造の導出を試みる．

3.2 コード化と適応度

ここでは，グラフ G を GA における 1 個体と考え，図 3 に示すように，G の隣接行列（上三角成分）を 1 列に並べたものを染色体とする [花田 07]．グラフ G を評価するための適応度 $F(G)$ は，前節の n4c-free MUG 構造の性質を満たすように，それぞれを次式のように表現する．

$$F(G) = \sum_{i=1}^{4} \omega_i \times f_i(G) \quad (1)$$

$$f_1(G) = \begin{cases} 1 & G \text{ が性質 1 を満たす} \\ 0 & Otherwise \end{cases} \quad (2)$$

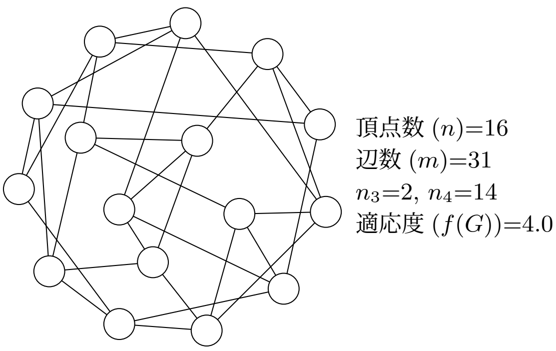


図 4: 導出された n4c-free MUG 構造の例

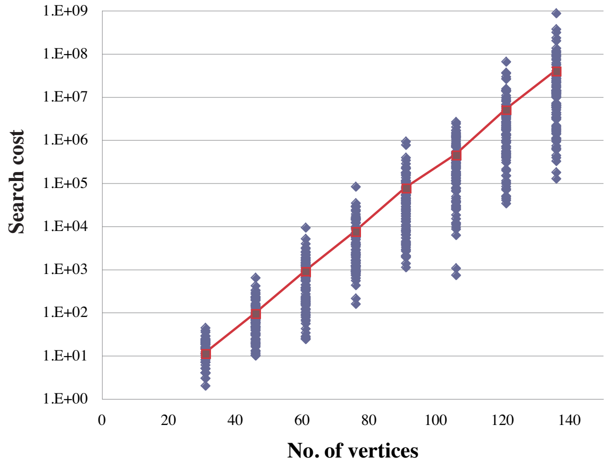


図 5: 実験結果（探索コスト）

$$
f_2(G) = \frac{n_3 + n_4}{n} \tag{3}
$$

$$
f_3(G) = \begin{cases} 1 & G \text{ が性質 3 を満たす} \\ 0 & Otherwise \end{cases} \tag{4}
$$

$$
f_4(G) = \begin{cases} 1 & var(G) \leq 1 \\ 1/var(G) & var(G) > 1 \end{cases} \tag{5}
$$

ここで，ω_i は関数 f_i に対する重み，n_j は次数 j の頂点数，$var(G)$ は G の頂点の次数の分散を表わしている．

3.3 導出結果

前節のコード化および適応度を用いて GA を実装し，文献 [Mizuno 08, Nagasawa 09] では扱っていない（つまり，試行錯誤では現実的な時間内に n4c-free MUG 構造を発見することができない）頂点数 n=16 の n4c-free MUG 構造の導出を試みた．ここでは，GA の世代交代モデルとして SGA，交叉方法として二点交叉を用いて，集団サイズ 20，世代数の上限は $2{\times}10^{12}$ とした．図 4 は，その結果導出された n=16 の n4c-free MUG 構造を表わしている．

4. 評価実験

4.1 実験条件と実験結果

本方法によって導出された n4c-free MUG 構造を用いて 3COL インスタンスを生成し実験を試みた．ここでは，図 4 の

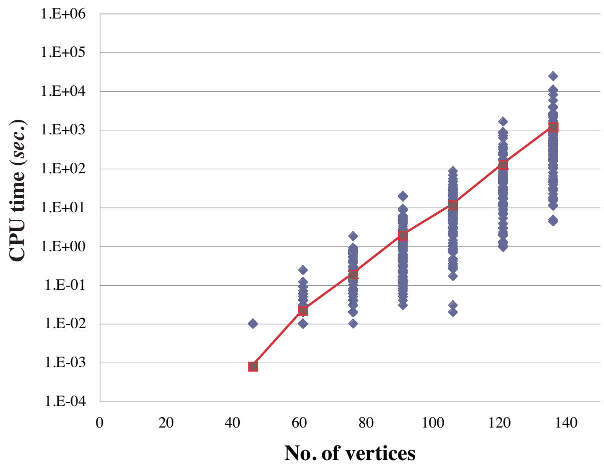


図 6: 実験結果（CPU time）

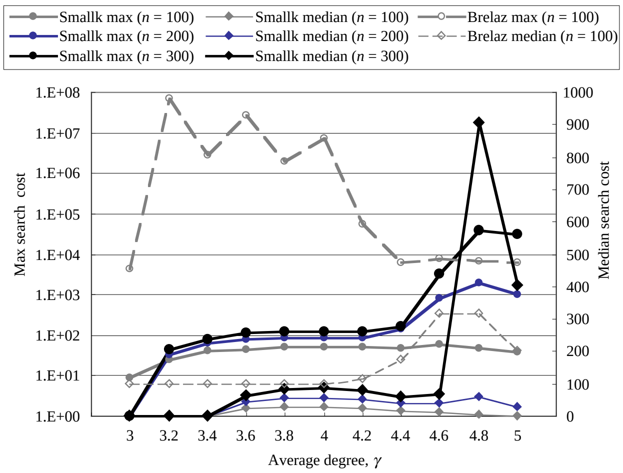


図 7: ランダムに生成したインスタンスの実験結果（文献 [Mizuno 08] より引用）

n4c-free MUG を図 2 の (1),(2) に与えることによって，インスタンスを生成した．埋め込み回数（図 2 のパラメータ k）を k=1∼8 の 8 ケース（頂点数は n=31∼136 となる）に対して，ケースごとに 100 個のインスタンスを生成した．また，生成したインスタンスを解くアルゴリズムとして，Smallk Coloring Program（以下，Smallk）[Culberson 01, Smallk 00] を用いて，アルゴリズム中の ‘Search node’（頂点への色の割当て回数に相当）を計数しそれを探索コストとした．

なお，実験には PC/AT 互換機（CPU: Intel Core 2 Quad Q9550 2.66GHz, RAM: 4096MB）を使用し，プログラムはすべて C 言語で記述した．

図 5，6 に，実験結果（それぞれ探索コストおよび探索時間）を示す．図 5 より，生成されたインスタンスのサイズが大きくなるにしたがって，探索コストがほぼ指数関数的に大きくなっていることが分かる．図 6 より，探索時間においても，図 5 にほぼ比例していることが分かる，これらの結果より，本方法で導出した n4c-free MUG 構造を用いて生成したインスタンスを解く計算量が，頂点数のほぼ指数オーダとなっていることが分かる．また，各インスタンスの計算量に注目すると，難しさ

のばらつきも小さく，最もコストや時間が小さいインスタンスにおいても，頂点数のほぼ指数オーダの計算量となっている．この結果は，文献 [Mizuno 08] で示されている実験結果とほぼ同等であり，したがって同程度の組合せ的複雑さを持つインスタンスが安定的に生成できているといえる．

また，図 7 は，文献 [Mizuno 08] より引用したランダムに生成したインスタンスに関する Brélaz ヒューリスティクスを組み込んだ木探索法（以下，Brélaz）[Brélaz 79]，Smallk の実験結果を表わしている．図 7 では，平均次数 γ を 11 ケースおよび頂点数 3 ケース，つまり 33 ケースに対して，各ケース 10,000 個のインスタンスを生成している（すなわち，全体では 330 万個）．このうち最も難しかったインスタンスの計算量を見てみると，Brélaz では頂点数 100 のケースで 1E+08 弱，Smallk では頂点数 300 のケースで 1E+07 強となっている．これを本研究によって生成されたインスタンスと比較すると，Brélaz，Smallk ともに，頂点数 80 付近でそれぞれの計算量を超えた手間がかかっており，膨大な数のランダムに生成されたインスタンスで最も難しいインスタンスよりも大きく手間がかかるインスタンスが生成できているということが分かる．

4.2 考察

4.1 節において，本研究で提案した生成手続きによるインスタンスを解く計算量が頂点数の指数オーダの手間になっていること，ランダムに生成したインスタンスを解く計算量より大幅に難しいインスタンスが生成できていること，およびそのような難しいインスタンスが安定的に生成できていることを確認した．これは，これらのインスタンスが，文献 [Mizuno 08] の生成手続きによるインスタンスと同様に，(1) インスタンス自体が極小非可解部分問題である，(2) 頂点の次数の分散が小さい，(3)n4c を部分問題として含まないので frozen pair が容易に発見できない，などの難しいインスタンスの諸性質をつねに満たしている結果であると考えられる．

つまり，本方法によって導出された n4c-free MUG 構造も，従来の試行錯誤で発見したものと同じ性質を持つことに加えて，それをもとに同等の難しさを持つ組合せ的複雑な 3COL インスタンスが生成可能であるといえる．したがって，本研究のような進化計算を利用した導出方法は，より多様な，かつより大きなサイズの n4c-free MUG 構造を発見することに対して有効である．しかしながら，今回導出した $n = 16$ の n4c-free MUG 構造を得るのに要する時間は，今回用いた計算機で数週間程度かかってしまう．これは試行錯誤的に導出を試みるよりははるかに高速であるものの，より大きなサイズの構造導出を行なうためには，GA の処理の高速化や適応度の改良などが必要である．ただし，n4c-free MUG 構造自体が，膨大なグラフ集合全体の中のほんの一部の領域にのみ存在しているためその発見が困難であるということも考えられるため，これらも踏まえてどの程度の高速化が図れるかは今後の重要な課題である．

5. おわりに

本報告では，解くのに手間のかかる難しい 3COL インスタンスを生成するために，重要な構造である n4c-free MUG 構造を GA によって導出する方法を提案した．また，新たな n4c-free MUG 構造を導出して，それを用いて生成されたインスタンスを解くことにより，その難しさが問題サイズのほぼ指数オーダとなることを実験で確認した．今後の課題としては，より大きなサイズの n4c-free MUG 構造導出のための，適応度関数の改良，および GA による処理の高速化などがあげられる．また，文献 [水野 10] で提案されているグラフの頂点の連結度に注目した生成方法に適用し，本研究で導出された n4c-free MUG 構造の有用性を確かめることが重要な課題である．

謝辞

本研究の一部は，科学研究費補助金（課題番号: 21700173）による．

参考文献

[Brélaz 79] Brélaz, D.: New methods to color the vertices of a graph, *Commun ACM*, Vol. 22, No. 4, pp. 251–256 (1979).

[Culberson 01] Culberson, J., Gent, I.: Frozen development in graph coloring, *Theoret. Comput. Sci.*, Vol. 265, pp. 227–264 (2001).

[花田 07] 花田良子, 佐藤史隆, 廣安知之, 三木光範, 鈴木泰博: 遺伝的アルゴリズムによるネットワーク特性量に着目したネットワーク設計法, コンピュータソフトウェア, Vol. 24, No.1, pp. 91–100 (2007).

[Hogg 94] Hogg, T., Williams, C. P.: The hardest constraint problem: a double phase transition, Artif. Intell., Vol. 69, pp. 359–377 (1994).

[Hogg 96] Hogg, T., Huberman, B.A., Williams, C.P.: Phase transition and search problem, *Artificial Intelligence*, Vol. 81, pp. 1–16 (1996).

[Mammen 97] Mammen, D. L., Hogg, T.: A New Look at Easy-Hard-Easy Pattern of Combinatorial Search Difficulty, *Journal of Artif. Intell. Res.*, Vol. 7, pp. 47–66 (1997).

[Mizuno 08] Mizuno, K, Nishihara, S.: Constructive generation of very hard 3-colorability instances, *Discrete Applied Mathematics*, Vol. 156(2), pp. 218–229 (2008).

[水野 10] 水野一徳, 長澤圭孝, 仲健次, 佐々木整, 西原清一: 極小非可解構造に基づく 連結度を考慮した 3COL インスタンスの組織的生成, 第 77 回人工知能基本問題研究会 (SIG-FPAI), 人工知能学会, pp. 67–72 (2010).

[Nagasawa 09] Nagasawa, Y., Mizuno, K., Sasaki, H., Miki, Y., Nishihara, S.: Constructive Generation of 3COL Instances by Embedding Minimal Unsolvable Structures, *Proc. KSE2009*, pp. 100–105 (2009).

[Smallk 00] Overview of the Smallk Graph Coloring Program (2000), http://www.cs.ualberta.ca/ joe/Coloring/smallk.html

[Vlasie 95] Vlasie, D. R.: Systematic Generation of Very Hard Cases for Graph3-Colorability, *Proc. the 7th IEEE ICTAI*, pp.114–119 (1995).