確率的制約充足アルゴリズムにおける 局所最適構造

Local Minimum Structures for Stochastic Constraint Satisfaction Algorithms

水野 一徳
Kazunori Mizuno

筑波大学 電子・情報工学系
Institute of Information Sciences and Electronics, University of Tsukuba
mizuno@algor.is.tsukuba.ac.jp, <http://www.npal.is.tsukuba.ac.jp/~mizuno/>

西原 清一
Seiichi Nishihara

(同 上)
nishihar@is.tsukuba.ac.jp, <http://www.npal.is.tsukuba.ac.jp/~nishihar/>

keywords: constraint satisfaction, search, hill-climbing, local minimum, stochastic search

Summary

Many stochastic search algorithms have recently been developed to make more remarkable progress than systematic search algorithms because stochastic algorithms sometimes solve large-scale constraint satisfaction problems in a practical time. However, such stochastic algorithms have the drawback of getting stuck in local optima which are not acceptable as final solutions.

We analyze an iterative improvement algorithm from the viewpoint of constraint structures causing local optima. Using the graph-coloring problem with three colors, an archetype problem to evaluate constraint satisfaction algorithms, we study the local graph structures around which so many conflicts thrash. We propose a key constraint structure, an LM pair, which may induce a local optimum, and clarify the mechanism of how conflicting colors for an LM pair obstruct the stepwise refinement of hill-climbing. Experimental results show that LM pairs are strongly correlated with the search efficiency of the stochastic search algorithm.

1. は じ め に

制約充足問題 (Constraint Satisfaction Problem: CSP) とは、離散値をとるいくつかの変数に割り当て可能な値の組合せのうち、与えられたすべての制約を満たすような組合せを探索によって発見する問題である [西原 97]。CSP は、設計や計画問題などを含む、人工知能やパターン処理など広い分野で応用されている基盤的技術である。CSP を解くために、従来よりバックトラックを基本とした木探索法に代表される系統的探索アルゴリズムが用いられてきた。系統的探索はアルゴリズムの完全性が保証されている反面、探索空間のサイズが大規模な CSP に対しては実用的な時間内に解探索を終了することが困難な場合がある。これに対して、近年、山登り法やニューラルネットワークなどを基本アルゴリズムとした反復改良型の探索アルゴリズムに関する研究が目覚ましい [Minton 92, Adorf 90, 水野 98]。本論文では、これを確率的探索アルゴリズムと呼ぶ。これは、アルゴリズムの完全性を犠牲にして、大規模な CSP に対して、実用的な時間内になるべく高い確率で解を発見する、ある

いは実用可能な近似解を高速に求めるというアプローチである。

しかし、確率的探索アルゴリズムには、解探索の途中で局所最適解に捕捉されてしまうという問題点がある。このため、局所最適解からいかに効率良く脱出するかといったメタ戦略に関する研究も盛んに行なわれており [Selman 92, Morris 93, 水野 98]、大規模な CSP に対して、比較的良好な結果が得られている。

CSP をはじめとする組合せ探索問題や組合せ最適化問題の多くは NP 完全であり、その意味で効率的な汎用解法は存在しないと考えられてきた。しかし、最近では、そのような組合せ的に複雑な具体的問題は意外に少なく、問題空間のある一部の領域のみに遍在しているということが明らかになっている [Mitchell 92]。また、そのような複雑な問題は、CSP の相転移 (Phase Transition: PT) 現象が起こる領域付近で発生することが観察されている [Cheeseman 91, Hogg 94, Hogg 96]。PT 領域付近の問題は、系統的探索アルゴリズムのみならず、確率的探索アルゴリズムを用いた場合でも解くのが困難であることが知られている。これに対して、文献 [Yokoo 97, 横尾

99] では、小規模な問題を用いて、問題の状態空間の地形を網羅的に解析することによって、確率的探索アルゴリズムでは、状態空間の中の解や解へ到達可能な状態および局所最適となる状態の数がその探索効率に大きな影響を与えていると報告されている。また、文献 [Frank 97] では、比較的大規模な問題に対して蓋然的な地形解析を行っている。

本論文では、CSP の確率的探索アルゴリズムの非効率の要因である局所最適解に陥ってしまう原因について、問題の制約グラフの構造に注目して解析する。具体的な対象問題としてグラフ色塗り問題、探索アルゴリズムとして反復山登り法を用いる。また、反復山登り法において、変数が持つ制約違反を解消することが不可能な状態について示し、局所最適となる状態を生成する可能性が高い制約グラフ構造である LM pair を提示する。さらに、LM pair が反復山登り法の解探索の効率に強く関連していることを示す。

本研究は確率的探索アルゴリズムにおける局所最適解が発生するメカニズムを解明する一試みであるが、このような解析を行なうことは、より効率的な制約充足アルゴリズムを開発するうえで重要であると考えられる。

以下、2 章では、まず、本論文の対象問題であるグラフ色塗り問題を定義する。次に、CSP の確率的探索アルゴリズムについて述べたあと、反復山登り法についてそのアルゴリズムを示して説明する。3 章では、グラフ色塗り問題について、局所最適となる可能性がある制約構造を提示する。4 章では、まず、グラフ色塗り問題を用いたシミュレーションにより反復山登り法の探索効率を評価した後、確率的探索アルゴリズムにおけるグラフ色塗り問題の難しさと制約構造の関係について考察する。

2. 研究分野の概要

2.1 制約充足問題 (グラフ色塗り問題)

本論文では、解析の対象としてグラフ色塗り問題 (Graph-Coloring Problem: GCP) を用いる。GCP とは、無向グラフの隣接頂点が同じ色にならないように、グラフのすべての頂点にある定められた個数の色のうちの 1 色を塗りわけていく問題である。GCP は NP 完全問題であり、種々の制約充足アルゴリズムの性能を評価するための標準的な問題として広く用いられている。また、問題の制約条件が単純かつ対称的であり、明示的にグラフの構造で表現することが容易であるという特徴を持つ。ここでは、GCP を 3 つ組 (V, C, E) で定義する (図 1 参照)*1。 $V = \{v_1, \dots, v_n\}$ は n 個の変数の集合で、各要素はグラフの頂点を表す。 $C = \{red, green, blue\}$ は変数に割り当てられる値の集合で、頂点に塗られる色の候補全体を表す。本論文では、色数 = 3 の GCP を対象とする。 $E = \{e_1, \dots, e_m\} (\subseteq V \times V)$ は m 個の 2 頂

$$GCP = (V, C, E)$$

$$V = \{v_1, \dots, v_n\}$$

$$C = \{red, green, blue\}$$

$$E = \{e_1, \dots, e_m\}$$

$$\text{ただし, } e = (v_p, v_q) \Leftrightarrow c_p \neq c_q$$

n : 変数 (頂点) の数

m : 制約 (辺) の数

c_i : 変数 v_i に割り当てられている値 (色)

図 1 グラフ色塗り問題

制約の集合を表しており、 $e(e \in E) = (v_p, v_q)$ は変数 v_p と v_q の間に制約が存在する、すなわち辺でつながれていることを示す。ここでは、 v_p と v_q が同じ色を割り当ててはいけないという制約条件がある。

また、本論文では、制約の数 m を変数の数 n で割ったものを制約密度 $d (= m/n)$ と定義し、この指標を用いて問題を分類する。GCP に関する研究では、制約密度 $d = 2 \sim 2.5$ 付近の領域で PT が起こることが観察されており [林元 99]、この領域における問題は系統的、確率的探索アルゴリズムいずれによっても解くのに手間のかかる難しい問題とされている。以上を踏まえて、本論文では、 $d = 1.6 \sim 6$ の GCP を対象とする。

2.2 確率的探索アルゴリズム

確率的探索アルゴリズムは、すべての変数に対して暫定的に値を割り当てた状態 (解候補) からスタートし、制約違反数を評価関数として、解候補が持つ制約違反が少なくなるように変数の値を修正することによって解探索を進めていく。これは、系統的探索アルゴリズムのように組合せの試行錯誤によって解探索を進めるのではなく、状態空間の地形を考慮しながら解の存在する可能性の高い状態への接近を試みる方法であるといえる。これによって、系統的探索アルゴリズムでは対処することが困難である大規模な CSP を解く場合でも、解を発見することが確率的に保証される。

しかし、解探索の途中で局所最適解に捕捉されてしまうことが問題点として指摘されている。このため、局所最適解に陥らないようにする、あるいは効率的に局所最適解から脱出するためのメタ戦略に関する研究が盛んに行なわれている [Selman 92, Morris 93, 水野 98]。また一方では、問題の状態空間の地形を解析するといった研究も行なわれており、解探索が非効率になる、すなわち解くのが困難な問題は、解に到達可能な状態の個数が減少するために局所最適解に陥りやすいということが報告されている [Frank 97, Yokoo 97, 横尾 99]。

確率的探索過程において現在の状態から次の状態へ遷

*1 (V, E) は、頂点集合 V 、辺集合 E の連結無向グラフ。

移させる操作は、ある 1 つの変数の値を異なる値に替える (repair) 操作である。局所最適状態とは、このような状態遷移操作によって制約違反数を減少させることができないような状態である。本論文では、解候補の状態が次のような場合、その状態は局所最適であると定義する。

- 解候補の 1 つの変数の値を異なる値に替えようとするとき、その変数に現在割り当てられている値以外の割り当て可能な値を割り当てたととしても、制約違反数が減少せず、このような状況が一定探索回数続いて、解候補の改善がみられない状態

2.3 反復山登り法

本論文では、CSP の標準的な確率的探索アルゴリズムとして用いられている、反復山登り法 (Iterated Hill-Climbing: IHC) を用いる。図 2 に IHC のアルゴリズムを示す。図 2 において、 s はすべての変数に何らかの値が割り当てられている解候補 (状態)、 $conf(s)$ は解候補 s の制約違反数を表している。IHC は、文献 [Minton 92] で述べられている制約違反最少化戦略による山登り法 (Min-Conflicts Hill-Climbing: MCHC) を基本探索アルゴリズムとしている。MCHC は、図 2 にも示すように、制約違反を持つ変数の中からランダムに変数を 1 つ選び、その変数に割り当て可能な値のうち、最も制約違反が少なくなる値を割り当てることにより探索を進める。また、MCHC による探索において、ある一定の探索回数 (山登りステップ数: 図 2 の max_check_LM) の間、制約違反が減少しなかった場合、その状態を局所最適とみなしてまた別の初期値から MCHC による探索をやり直すというものである。

3. 局所最適状態を惹き起こす制約構造

2.2 節において、IHC をはじめとする CSP の確率的探索アルゴリズムの解探索が非効率になるときの原因の 1 つとして、局所最適解の存在が影響しているということについて説明した。それでは、2.3 節で説明した IHC を用いて GCP を解くとき、解候補の状態がどのような場合に局所最適となってしまうのだろうか？ここで、図 3 のような状態を考える。図 3 において、頂点 A の次数は δ 、すなわち、隣接している頂点が δ 個あるものとする。ここで、 δ 個の隣接頂点に塗られている色のうち、赤 (*red*)、緑 (*green*)、青 (*blue*) が塗られている頂点の個数を、それぞれ r, g, b (≥ 1) 個とすると、 $\delta = r + g + b$ であり、頂点 A は r 個の制約違反に参加している。ここで、頂点 A と同じ色の隣接頂点の個数が、それ以外のいずれよりも少ない場合を考える。すなわち、

$$b > r \geq 1 \quad \text{かつ} \quad g > r \geq 1 \quad \dots (\text{条件} *)$$

なる条件が成り立つ場合を考える。

図 2 に示すように、IHC は制約違反を持つ変数の中からランダムに変数を 1 つ選択して (変数選択処理)、その

```

procedure IHC (max_iteration,
                max_hc_step, max_check_LM)
begin
  for  $i := 1$  to max_iteration do
     $s :=$  random valuation of variables;
    check_LM := 0;
  /* Min-Conflicts Hill-Climbing */
  for  $j := 1$  to max_hc_step do
    if  $conf(s) = 0$  then return  $s$ ;
    choose randomly a variable  $v$  in conflicts;
    assign a value  $c$  that minimizes the number of
      conflicts for  $v$  : ( $s \rightarrow s'$ );
    if  $conf(s') < conf(s)$  then
      check_LM := 0;
    else
      check_LM := check_LM + 1;
      if check_LM = max_check_LM then break;
    end if
  end for
end for
end IHC

```

図 2 IHC のアルゴリズム

変数に割り当て可能な値のうち最も制約違反数が少なくなるような値を割り当て (値選択処理) ことにより解探索を進めていく。ここで、図 3 の状態において、IHC が変数選択の処理において頂点 A を選択したとする。このとき、頂点 A に割り当て可能な値として現在割り当てられている *red* 以外に *green*, *blue* があるが、*green*, *blue* を頂点 A に割り当てようとすると、いずれの値も制約違反数が増えてしまうため割り当てることができない。すなわち、この場合頂点 A は *red* 以外の色を塗ることができず、IHC では頂点 A が関与する制約違反を直接的に充足することができない。

ここで (条件 *) が成り立つような値が割り当てられる可能性がある頂点同士が辺でつながれている (制約が存在する) とき、その頂点組を “LM (local minimum) pair” と呼ぶ。以下に LM pair を定義する。

【定義 1】(LM pair) 色数 k の GCP において、次数 $2k - 1$ 以上の隣接頂点組。

すなわち、LM pair とは、局所最適となる状態を生成する可能性のある構造をした頂点組であり、LM pair のそれぞれの頂点への値の割り当てが (条件 *) と同様の条件を満たし、かつその LM pair に同じ色が塗られている場合、その状態は局所最適となってしまう。

図 4 は、LM pair (頂点組 B-C) を含んでいる部分グラフ構造の例を示している。図 4 において、頂点 B, C

図3 制約違反が解消できない頂点

図4 最も単純な LM subgraph の例

をみるといずれも $\delta = 5$, $r = 1$, $g = b = 2$ となり, 図3における頂点Aと同じ(条件*)を満たす. また, LM pair 間には制約違反があるので, IHC を用いて頂点B-C間の制約違反を解消することは不可能である. すなわち, IHC では, この状態を破棄して別の初期値から解探索をやり直す(リスタート)しかないといえる. 本論文では, このように LM pair とそれに隣接する頂点によって構成される誘導部分グラフ(induced subgraph)の構造を“LM subgraph”と呼ぶ. 実は, 図4は, LM subgraphのうち単純な部分グラフの構造の一例であり, 同時に局所最適となってしまう変数への値の割り当てを示していることになる.

IHCの探索過程において局所最適に陥ったときのいくつかのケースについて, それらの状態における制約違反の状況を観察したところ, そのほとんどが LM pair 間の制約違反であることが分かった. したがって, このような制約が問題(グラフ)中に多く存在するとき(条件*)を満たしてしまう可能性が高くなると考えられる. つまり, 局所最適となる状態に陥る可能性が高くなり, その

図5 図4における LM pair 間の制約違反が LM cycle によって解消できた例

問題は解くのが難しくなると考えられる.

制約密度が低いとき, すなわち制約の数が少ないとき, 各頂点の次数は低いために問題中に含まれる LM subgraph は少なく, それらの問題の解は容易に見えてくる. では制約を増やしていったとき, 問題中に含まれる LM subgraph の数は増えていき, すると解探索は制約密度が高くなるほど非効率になっていくはずである. しかし, 文献[水野98]や[横尾99]では, 系統的探索アルゴリズムを用いて観察されている相転移現象と類似した現象を, 確率的探索アルゴリズムにおいても観察している. つまり, 解探索コスト(問題の難しさ)が, 制約密度に対して, easy-hard-easy のコストパターン[Mammen 97]を形成し, 制約密度が高くなっていき, ある領域(hard)を越えたとき, 解探索は迅速(easy)になる. したがって, 制約密度が高くなっていったとき, LM subgraph 内にも制約が増えていくことにより, LM pair 間の制約違反を解消できる状態になっていくものと考えられる.

ここで, 図4における頂点DとEとの間に制約が存在する場合を考える. このとき, IHCではLM pairの制約違反を充足できる可能性が出てくる. これは, 修正不可能な LM pair 間の制約違反以外に, 頂点D-E間にも制約違反が発生することにより, IHCがこの頂点D-E間の制約違反を修正しようとする. 例えば, 頂点Eにgreenを割り当てると制約違反が減少し, 頂点Bには, blueも割り当て可能になる. 次にIHCがランダムに頂点Bを選択してblueを割り当てたとすると, この時点で頂点B-D間の制約違反のみとなる. さらに, その次の山登りステップで頂点Dが選択されると, 必然的に頂点Dにはredが割り当てられることになる. 最終的に図5のような割り当てになり, LM pair 間の制約違反は解消される. 図4の頂点D-E間の制約が存在するという構造は, LM subgraph中に頂点B, D, Eの3頂点からなるサイクル(3-cycle)が1つ含まれている構造とみることができる.

ここで“LM cycle”を導入しておく.

【定義2】(LM cycle) LM subgraphの部分グラフであって, かつ LM pairの頂点が1つ参加している3-cycle.

LM cycle は, LM pair の制約違反を充足できるようにして, その結果として局所最適状態を減少させる. したがって, 問題に LM subgraph が多く含まれている場合, 局所最適状態に陥りやすくなり問題は難しくなるが, LM cycle の総数が多くなるほど, 局所最適解またはそれに至る状態が解またはそれに至る状態に変化していく. 結果としてその問題は解くのが容易になると考えられる.

4. 評価

4.1 IHC の探索効率 (GCP の難しさ)

GCP の難しさを評価するため, 頂点数 $n = 50, 80, 100$, 色数 = 3 の GCP に対して, IHC を用いて解探索を試みた. これらの問題における探索空間のサイズはそれぞれ $3^{50} (\approx 10^{24})$, $3^{80} (\approx 10^{38})$, $3^{100} (\approx 10^{47})$ となる. また, ここでは制約密度 $1.6 \leq d \leq 6$ について, $d = 1.6, 1.8, 2, 2.2, 2.3, 2.4, 2.5, 2.6, 2.8, 3, 4, 5, 6$ の計 13 ケースの問題を対象とした. 問題は文献 [Minton 92] で示されている方法で, 解が存在する (可解な) 問題を疑似乱数を用いて生成し, その中からグラフが連結しているもののみについて, 各ケースごとに 100 問 (計 1,300 問 $\times 3$) の GCP を用意した. IHC において, 探索開始時およびリスタート時における初期値は, 疑似乱数を用いてランダムに生成するものとした. また, 図 2 における $max\_iteration$, max_hc_step , max_check_LM は, それぞれ 100, 1,000, 100 に設定した. よって, 山登りステップ総数の上限すなわち制限探索回数は, $max\_iteration \times max\_hc\_step = 10^5$ となる.

効率評価は次の 2 つの項目について行なった.

(1) 探索成功率 (%) :

探索した GCP のうち, 制限探索回数内に解を発見することができたものの割合

(2) 平均探索時間 (sec) :

解を発見することができた問題についての探索時間の平均値

なお, 実験は PC/AT 互換機 (CPU: Celeron 400 MHz, メモリ: 192 Mbyte) を使用し, プログラムは C 言語で記述した.

図 6 に実験結果を示す. 制約密度 $d = 2.4$ 付近の領域において, 探索成功率が大きく悪化していることが分かる. また, この領域に近づくにしたがって, 解を発見するためにより多くの探索時間を要していることが分かる. これは, GCP の相転移現象による解探索の非効率が起こる領域とほとんど一致している. よって, この領域付近の問題は, IHC では局所最適解に陥りやすくなることにより, 解くのが難しくなっているものと考えられる. さらに, 頂点数 n が大きくなる, すなわち探索空間のサイズが大きくなるにしたがって解探索が非効率になっていくことが分かる.

図 6 IHC の探索効率

図 7 LM pair の数と 3-cycle の数の割合

4.2 LM subgraph の解析

本節では, 4.1 節の実験で用いたのと同じ GCP 集合について, 各問題に含まれている LM pair の数と 3-cycle の数を解析した. 図 7 にその解析結果を示す. ここで, 割合 α は, 問題に含まれていた LM pair の数を 3-cycle の数で割ったもの (LM pair の数 / 3-cycle の数) とし, 図 7 の結果は, 制約密度の各ケースごとにおける 100 問に対する割合 α の平均値を表している.

図 6 と 図 7 を比較すると, 頂点数 n のいずれの場合においても図 7 の割合 α が図 6 における平均探索時間の結果と類似したパターンが得られており, 解探索が非効率であった制約密度 $d = 2.4$ 付近において割合 α が最も高い値を示していることが分かる.

また, 図 8 は, 図 7 のうち $n = 100$ における制約密度 13 ケースの問題 (1,300 問) について, IHC の探索効

図 8 それぞれの α に基づく IHC の探索効率の分類

表 1 制約数に対する LM pair 数の割合と平均 LM cycle 数

制約密度 (d)	1.6	1.8	2.0	2.2	2.3	2.4	2.5	2.6	2.8	3.0	...	4.0	...	5.0	...	6.0
制約数に対する LM pair 数の割合 (β)	0.12	0.21	0.30	0.39	0.44	0.49	0.53	0.58	0.65	0.73	...	0.93	...	0.99	...	1.00
平均 LM cycle 数 (γ)	0.71	0.87	1.12	1.35	1.40	1.49	1.66	1.86	2.14	2.39	...	4.79	...	8.80	...	14.74
比 β / γ	0.17	0.24	0.27	0.29	0.31	0.33	0.32	0.31	0.30	0.31	...	0.19	...	0.11	...	0.07

率と各問題の α との関係について分類したものである。図 8 において、解を発見するに至らなかった問題 (not solved) は、制限探索回数が終了した時点での探索時間をプロットした^{*2}。また、図 8 中の曲線は、 $1 \leq \alpha < 20$ において各 $\alpha \pm 1$ 区間の問題における探索時間の平均値を表している。図 8 より、制限探索回数内に解を発見することができた (solved) 問題に比べて、not solved の問題は比較的 α が高い領域に集中していることが分かる。平均値の曲線を見ると、 $\alpha = 11$ 付近までは、ほぼ線形的に増加していることが分かる。 $\alpha = 11$ 付近以降の平均値は幾分減少傾向にあるが、not solved の問題がより多くの探索時間を要することを考えると、この領域以降も線形的に増加する傾向にあると考えられる。なお、not solved の問題に関する α の平均値は 9.15、また、 α と探索時間の相関係数は、0.469 であった。

次に、 $n = 100$ の各問題に対して、制約の数に対する LM pair の数の割合 β ($=$ LM pair の数 / 制約数) と各 LM subgraph に平均いくつの LM cycle が含まれていたか (平均 LM cycle 数: γ) について調べた。表 1 にその結果を示す。表 1 におけるそれぞれの値は、制約密度の各ケースごとの 100 問の平均値である。表 1 より、まず、制約密度が高くなる、すなわち制約が増加していくと、割合 β 、平均 LM cycle 数 γ ともに増加していく。しかし、制約密度が低い領域から制約密度 $d = 2.4$ 付近の領域にかけては、割合 β が急激に増加しているのに対して、平均 LM cycle 数 γ はほとんど増加せず、ほぼ一定の値を示している。また、制約密度 $d = 2.4$ 以降は制約密度が高くなっていくと、 β は徐々に収束しているのに対して、 γ が急激に増加しはじめていることが分かる。

さらに、 β/γ は、図 7 と同様に、制約密度 $d = 2.4$ に

近づくにしたがってこの値が高くなっていき、 $d = 2.4$ で最も高い値を示していることが分かる。

4.3 考 察

4.1 節の実験結果と 4.2 節の解析結果より、LM subgraph と、解探索の効率、すなわち問題の難しさとの関係について考察する。

まず、制約密度が低い領域では、問題中のすべての制約のうち、LM pair 間の制約となっているものの割合が非常に少ない。したがって、この領域における問題は、局所最適状態に陥る可能性が極めて低く、解の個数も多いために、解くのが容易な問題となり解探索は効率的に行なうことができるものと考えられる。

しかし、制約が徐々に増えていき、制約密度 $d = 2.4$ 付近の領域に近づくにつれて、問題の制約のうちの LM pair の数の割合が急激に増加する。それに伴い、局所最適状態に陥ってしまう可能性が高くなることで問題は難しくなり、一定時間内に解を発見することが困難になる。

さらに、制約が増えていくと、問題の制約はほぼすべてが LM pair 間の制約となる。しかし、制約密度 $d = 2.4$ 付近の領域以降、問題中の 3-cycle 数が急激に増加してくるために、各 LM subgraph は、より多くの LM cycle を持つようになる。LM cycle は、LM subgraph が生成する局所最適状態を減少させるため、制約密度が高い領域の問題では、局所最適状態数は少ないと考えられる。このため、迅速な解探索が行なえるものとする。

ところで、任意の 2 頂点が LM pair かどうかは、GCP の場合、問題中のすべての制約の頂点組についてチェックすれば判定できる。よって、その判定に要する計算量は、制約の数が m なら $O(m)$ である。したがって、これを解探索の前処理として問題中の制約が LM pair かど

*2 実際に解を発見するには、これより多くの時間を要する。

図 9 LM pair 構造以外で局所最適を生成する構造の例

うか判定した後，LM pair を解探索の初期値生成に利用するか，あるいは LM pair の制約に重み付けを行うなどして解探索の効率化が図れる可能性があると考えられる．

最後に，GCP におけるすべての局所最適解は，LM pair 構造によって発生するののかという問題がある．GCP におけるほとんどの局所最適は，図 3 の頂点 A の（条件 *）下にある頂点同士が辺でつながれているか（LM pair），あるいは，頂点 A の条件下にある 2 頂点の間に別の頂点がいくつか接続しているものである．しかし，図 9 のように 図 3 の頂点 A の条件を満たしている頂点が存在しないにもかかわらず，局所最適な状態を生成する構造が存在する．図 9 では，頂点 F, G, H の間で制約違反が巡回して，解の状態に到達できない構造を表している．しかし，このような構造は制約密度が高くなるにしたがって，ほとんど問題中に現れなくなるものと考えられるが，今後はこのような構造も含めて解析を行なうことは興味深い課題であると考えられる．

また，文献 [Yokoo 97, 横尾 99] では，局所最適解の個数は，制約が増えるにしたがって，単調減少する傾向にあり，PT 領域付近では，解の個数がそれよりも急激に減少するために解に到達可能な状態が減少することで問題が難しくなるということを指摘している．本論文における構造解析では，IHC が局所最適に陥りやすい構造を示すことができたが，PT 現象との直接の関連性について明らかにしたわけではない．今後は，問題の状態空間の地形解析を行ない，解や解に到達可能な状態，および局所最適となる状態の個数などと，LM pair 構造との相関関係について考察することによって PT 現象との関連を明らかにすることが重要である．

5. お わ り に

本論文では，制約充足問題の確率的探索アルゴリズムの非効率の要因である局所最適解に捕捉される原因について解析した．グラフ色塗り問題に対して，反復山登り法が制約違反を解消することができない構造を示し，局所最適となる状態が生成される可能性が高い制約構造（LM pair）と LM pair を含んでいる部分グラフ構造（LM

subgraph）を提示した．また，LM subgraph が生成する局所最適状態に対して，制約違反が解消できるようになる状況（LM cycle）を示した．さらに，問題により多くの LM subgraph が含まれている場合，反復山登り法の解探索は非効率となる（困難な問題）．しかし，より多くの LM cycle が存在する領域の問題は解くのが容易であること，すなわち解探索が迅速に行なえることを確認した．

本論文で提示した LM pair 構造は，グラフ色塗り問題が持つ特有の制約構造によって，明示的に部分グラフ構造として定義することができた．しかし，一般の CSP では，より複雑な制約条件を持つ問題が多く，このように容易に明示的構造として表すことは困難である．しかし，一般の CSP においても制約違反が解消できない状況があり，それによって局所最適解が生成されているものと考えられる．したがって，今後は，他の CSP や WSAT [Selman 94] などの他の探索アルゴリズムを用いて詳細に解析することによって，統合的な局所最適解の発生メカニズムを解明することは重要であると考えられる．さらに，それらの結果に基づいて，より効率的な制約充足アルゴリズムを開発することは重要な課題であると考えられる．

謝 辞

本研究を進めるにあたって有益なご意見をいただいた，本学電子・情報工学系 福井幸男教授，本学工学研究科 小野智司氏，本学理工学研究科 林元充輝氏に深く感謝いたします．

◇ 参 考 文 献 ◇

- [Adorf 90] Adorf, H. M. and Johnston, M. D.: A Discrete Stochastic Neural Network Algorithm for Constraint Satisfaction Problems, in *Proceedings of IJCNN'90*, pp. 917-924 (1990).
- [Cheeseman 91] Cheeseman, P., Kanefsky, B., and Taylor, W.: Where the Really Hard Problems Are, in *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI'91)*, pp. 331-337 (1991).
- [Frank 97] Frank, J., Cheeseman, P., and Stutz, J.: When Gravity Fails: Local Search Topology, *Journal of Artificial Intelligence Research*, Vol. 7, pp. 249-281 (1997).
- [林元 99] 林元 充輝, 水野 一徳, 西原 清一: グラフ色塗り問題の制約構造に基づく探索効率の解析, 第 39 回人工知能基礎論研究会 (SIG-FAI-9903-1), pp. 1-6, 人工知能学会 (1999).
- [Hertz 94] Hertz, A., Jaumard, B., and Aragão, M. P.: Local optima topology for the k -coloring problem, *Discrete Applied Mathematics*, Vol. 49, pp. 257-280 (1994).
- [Hogg 94] Hogg, T. and Williams, C. P.: The hardest constraint problems: a double phase transition, *Artificial Intelligence*, Vol. 69, pp. 359-377 (1994).
- [Hogg 96] Hogg, T., Huberman, B. A., and Williams, C. P.: Phase transition and search problem, *Artificial Intelligence*, Vol. 81, pp. 1-15 (1996).
- [Mammen 97] Mammen, D. L. and Hogg, T.: A New Look at the Easy-Hard-Easy Pattern of Combinatorial Search Difficulty, *Journal of Artificial Intelligence Research*, Vol. 7, pp. 47-66 (1997).
- [Minton 92] Minton, S., Johnston, M. D., Philips, A. B., and Laird, P.: Minimizing conflicts: a heuristic repair method

for constraint satisfaction and scheduling problem, *Artificial Intelligence*, Vol. 58, pp. 161–205 (1992).

[Mitchell 92] Mitchell, D., Selman, B., and Levesque, H.: Hard and Easy Distributions of SAT Problems, in *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI'92)*, pp. 459–465 (1992).

[水野 98] 水野 一徳, 狩野 均, 西原 清一: 適応型確率探索による制約充足問題の解法, 情報処理学会論文誌, Vol. 39, No. 8, pp. 2413–2420 (1998).

[Mizuno 00] Mizuno, K. and Nishihara, S.: Local Minimum Structures of Graph-Coloring Problems for Stochastic Constraint Satisfaction Algorithms, in *The Twelfth IEEE International Conference on Tools with Artificial Intelligence (ICTAI'2000)* (2000).

[水野 00] 水野 一徳, 西原 清一: 制約充足問題の確率的解探索における局所最適の構造に基づく解析, 情報処理学会第 60 回全国大会, 3H-2 (2000).

[Morris 93] Morris, P.: The Breakout Method for Escaping From Local Minima, in *Proceedings of the Eleventh National Conference on Artificial Intelligence (AAAI'93)*, pp. 40–45 (1993).

[西原 97] 西原 清一: 制約充足問題の基礎と展望, 人工知能学会誌, Vol. 12, No. 3, pp. 351–358 (1997).

[Selman 92] Selman, B., Levesque, H., and Mitchell, D.: A New Method for Solving Hard Satisfiability Problems, in *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI'92)*, pp. 440–446 (1992).

[Selman 94] Selman, B., Kautz, H., and Cohen, B.: Noise Strategies for Improving Local Search, in *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI'94)*, pp. 337–343 (1994).

[Yokoo 97] Yokoo, M.: Why Adding More Constraints Makes a Problem Easier for Hill-Climbing Algorithms: Analyzing Landscape of CSPs, in *Proceedings of the Third International Conference on Principles and Practice of Constraint Programming (CP'97)*, pp. 356–370 (1997).

[横尾 99] 横尾 真: 制約充足問題の地形の解析, コンピュータソフトウェア, Vol. 16, No. 3, pp. 57–65 (1999).

〔担当委員: 阿久津達也〕

2000 年 5 月 11 日 受理

図 A.1 IHC の探索効率と α ($n = 100$, 色数 = 4 の GCP)

◇ 付 録 ◇

A. 色数 4 の GCP に関する評価

4 章で用いた頂点数 $n = 100$ の GCP に対して, 色数を 4 にしたときの IHC の探索効率と各問題の α について解析した. ここでは制約密度 $d = 2, 2.5, 3, 3.5, 4, 4.5, 5$ の計 7 ケースの問題を対象とし, 各ケースごとに 100 問の問題をランダムに生成した. また, LM pair は 3 章の定義より, 次数 $7 (= 2 \times 4 - 1)$ 以上の隣接する頂点組として, その個数をカウントした. 図 A.1 に実験結果を示す. ここで, IHC は 4.1 節と同様の条件で解探索を行なった結果, 制約密度のすべてのケースにおけるすべての問題について制限探索回数内に解を発見することができた. 図 A.1 より, 制約密度 $d = 4$ 付近において, 解を発見するまでにより多くの探索時間を要することが分かる. また, α についてもほぼ同じ領域において最も高い値を示していることが分かる.

著 者 紹 介

水野 一徳(学生会員)

1996 年筑波大学第三学群情報学類卒業. 1998 年同大学院理工学研究科修士課程修了. 現在, 同大学院工学研究科博士課程在学中. 知識処理, 制約充足問題, 組合せ探索アルゴリズムの研究に従事. 情報処理学会会員.

西原 清一(正会員)

1968 年京都大学数理工学科卒業. 同年, 同大学大型計算機センター助手. 1975 年より筑波大学電子・情報工学系. 現在, 同教授. 工学博士. 1982–1983 年ヴァージニア工科大学, 1998 年 IASA. グラフィクスと CAD, 組合せ探索アルゴリズム, 知識処理, 制約充足問題, 複雑系の研究に従事. 著書に「データ構造」(オーム社)等. 1975 年情報処理学会論文賞. 情報処理学会, 電子情報通信学会, ACM, IEEE 各会員.