1M2-OS-02b-1

研究室配属問題のCSP符号化手法の検討

Investigation of CSP Encoding approach for laboratory assignment problems

藤井 樹 [*1]　伊藤 靖展 [*2]　鍋島 英知 [*3]
Tatsuki fujii　Yasunobu itou　Hidetomo Nabesima

[*1]山梨大学工学部コンピュータ理工学科
Department of Computer Science and Engineering, Faculty of Enginnering, University of Yamanashi

[*2]旧山梨大学工学部コンピュータ理工学科
Former Department of Computer Science and Engineering, Faculty of Enginnering, University of Yamanashi

[*3]山梨大学大学院医学工学総合研究部
Department of Research Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi

We consider the laboratory assignment problem in which laboratories have minimum and maximum quotas. MSDA proposed by Ueda et al., is an efficient algorithm to solve the laboratory assignment problem, but it is an incomplete to find a fair assignment. In this paper, we propose a CSP encoding approach for the laboratory assignment problem which can find a fair assignment if it exist, and allows partial order on student's preference of laboratries. The experimental results show that the partial order make it possible to find more desirable assignments for students.

1. はじめに

研究室配属問題とは，研究室と学生の集合が与えられたときに，学生達にとって望ましい割当てを求める問題である．研究室が配属人数に対して個別の上限と下限を持つときに, 配属の割当てを効率良く求めるアルゴリズム MSDA(Mutli Stage Deferred Acceptance)[1] が Ueda らによって提案されている. MSDA は学生が妥当な不満を持たないフェアな解を求めるアルゴリズムであるが, たとえフェアな解が存在してもそれを検出できない場合がある (この場合 MSDA は, マスターリスト (Master List; ML) と呼ばれる学生の順位付けに基づく ML フェアな解を出力する).

本論文では, 研究室配属問題を制約充足問題 (Constraint Satisfaction Problem; CSP) に符号化して解く手法を提案する．本手法では, フェアな解が存在する場合には必ずそれを求めることが可能である．また, 問題によってはフェアな解が無数に存在することがあるが, 本手法ではそれらを学生の満足度に基づいて順序付けし, 満足度が最大となる解を求めることもできる．さらに MSDA では各学生の研究室に対する選好順位は全順序とする必要があったが, 本論文ではその制約を緩め, 半順序を認める手法も示す．これにより, 研究室配属問題がより現実的になるとともに, CSP 上の変数ドメイン小さく抑えることが可能になり, 求解を高速化することができる．

本手法は MSDA よりも低速ではあるが, 学生にとってより望ましい解を発見可能である．また学生の研究室に対する選好を全順序から半順序に緩めることで, 各学生がより上位に希望する研究室に配属される可能性が高くなることを評価実験によって実証する．

本論文の構成は以下の通りである．まず 2 章で研究室配属問題についての詳細を説明する．3 章では関連研究について述べる．4 章で今回の提案手法である CSP 符号化について説明, 5 章で同順位を含んだ場合の CSP 符号化について説明する．6 章では提案手法の実験による評価を行う．7 章ではまとめと今後の課題を述べる．

連絡先: 藤井 樹，山梨大学工学部コンピュータ理工学科，〒 400-8511 山梨県甲府市武田 4-3-11，E-mail : fujii@nabelab.org

2. 研究室配属問題

研究室配属問題は, 研究室と学生の集合が与えられたときに, 学生たちにとって望ましい割当てを求める問題であり, 以下の要素から構成される．学生の集合を S, 研究室の集合 L とし, 各学生 $s \in S$ の研究室に対する選好順位を $\succ_s$, 各研究室 $l \in L$ の学生に対する選好順位を $\succ_l$ で表す．また各研究室 $l \in L$ は受け入れ可能な人数の上限 ub_l と下限 lb_l をもつ．ここで学生数は下限の合計以上, 上限の合計以下であると仮定する．すなわち $\sum_{l \in L} \mathrm{lb}_l \leq |S| \leq \sum_{l \in L} \mathrm{ub}_l$ である．またすべての学生を GPA 等をもとにして順位付けしたマスターリスト (Master List; ML) $\succ_{ML}$ があると仮定する．各学生の研究室への割当ては写像 $M : S \to L$ により表す．また研究室に割り当てられた学生集合も同様に写像 $M : L \to 2^S$ により表すものとする．次節以降では, 研究室配属問題において望ましい割当てが満たすべき条件を紹介する．

2.1 無駄がない

ある割当て M において学生 $s \in S$ が, 配属する研究室 l よりも, 行きたい研究室 $l' \in L$ が存在したとする ($l' \succ_s l$). このとき, 学生 s が研究室 l から l' へ移動したとしても, l の下限と l' の上限の制約を満たすならば, 学生 s は研究室 l' に**空きシートを要求する**という．形式的には下記の 3 つの条件を満たす場合, 学生 s は研究室 l' に空きシートを要求する．

$$\exists\, l, l' \in L \begin{pmatrix} l' \succ_s l \\ \mid M(l') \mid < \mathrm{ub}_{l'} \\ \mid M(l) \mid > \mathrm{lb}_l \end{pmatrix}$$

空きシートを要求する学生がいない割当てのことを**無駄がない**という．

2.2 フェア

ある割当て M において学生 $s \in S$ が, 配属する研究室 l よりも, 行きたい研究室 $l' \in L$ が存在したとする $(l' \succ_s l)$. さらに l' に割り当てられた学生 $s' \in S$ について, 研究室 l' の選好順位において s のほうが s' よりも高いとき $(s \succ_{l'} s')$, 学生 s は学生 s' に**妥当な不満**を持つ. 形式的には下記の 3 つの条件をを満たす場合, 学生 s は学生 s' に妥当な不満を持つ.

$$\exists\, l, l' \in L, \exists\, s' \in S \begin{pmatrix} l' \succ_s l \\ s' \in M(l') \\ s \succ_{l'} s' \end{pmatrix}$$

妥当な不満を持つ学生がいない割当てのことを**フェア**という.

研究室配属問題において研究室に上限・下限制約が存在する場合, フェアかつ無駄がない割当てが存在しない場合があることが知られている [2]. そこで, フェアかつ無駄がない割当てがない場合でも, 任意の学生に対して, より行きたい研究室において, その研究室の選好順位に加えて, ML 上でも下位の学生がいないことを保証するフェアよりも条件の緩い ML フェアが存在する.

2.3 ML フェア

ある割当て M において学生 $s \in S$ が, 配属する研究室 l よりも, 行きたい研究室 $l' \in L$ が存在したとする $(l' \succ_s l)$. さらに l' に割り当てられた学生 $s' \in S$ について, 研究室 l' の選好順位において s のほうが s' よりも高く $(s \succ_{l'} s')$, ML 上でも s のほうが上位のとき $(s \succ_{ML} s')$, 学生 s は学生 s' に**強い妥当な不満**を持つ. 形式的には下記の 4 つの条件を満たす場合, 学生 s は学生 s' に強い妥当な不満を持つ.

$$\exists\, l \in L, \exists\, s' \in S \begin{pmatrix} l' \succ_s l \\ s' \in M(l') \\ s \succ_{l'} s' \\ s \succ_{ML} s' \end{pmatrix}$$

強い妥当な不満を持つ学生がいない割当てのことを **ML フェア**という.

3. 関連研究

研究室配属問題を解く既存の研究として上田らは, MSDA[1] を提案した. MSDA は DA アルゴリズム [3] を拡張したアルゴリズムであり, DA アルゴリズムでは上限制約しか扱うことができなかったため, それを下限制約が存在する場合でも使用できるように拡張した. MSDA は, 下限の合計分の学生を常時確保しつつ仮配属を行い, 下限の合計と確保される学生の数が同数になったら, 残りの学生を下限を満たすように配属させていけば, 必ず下限が満たされる割当てが求めれられるという手法である. MSDA で求める割り当ては, 無駄がないかつ ML フェアであることが保証されている.

4. 研究室配属問題の CSP 符号化

本論文では, 研究室配属問題を整数有限領域上の CSP (Constraint Satisfaction Problem) へと符号化を行うことで解を求解する手法を提案する. CSP とは, 制約充足問題のことで, 変数に与えられたドメインから値を割り当てることで, 与えられた制約すべてを満たすことができる解が存在するか判定する問題である. 整数有限領域上の制約充足問題は, 形式的には以下のように定義できる [4].

制約充足問題は以下を満たす組 $\mathrm{CSP}(X, Dom, C)$ である. 制約には, 算術論理演算等で条件が記述されている内包的制約, 制約を満足する (あるいは制約に違反する) 値の組の集合が陽に与えられている外延的制約, そして alldifferent 等に代表されるいわゆるグローバル制約がある.

CSP ソルバーは, 制約すべてを満たすような変数の割り当てが存在するのであれば充足可能を返し, 割り当てを出力する. これを利用し, "無駄がない"や"フェア"などを制約として CSP に変換することで, フェアかつ無駄がない割当てを探索する. 図 1 に CSP 符号化手法の手順を示す. 研究室配属問題を制約として表現して CSP へと符号化を行う. 符号化された問題を CSP ソルバーに解かせることで CSP の解を得て, それを復号化することで元問題の解を得る. フェアな解が存在しないときは ML フェアを制約化して探索する. 下記及び次節以降に割当ての望ましさを表す制約を紹介する.

ある学生 $s \in S$ がある研究室 $l \in L$ に配属されているかどうかを表す変数として, $Assign(s, l)$ を用いる. この変数のドメインを $\{0, 1\}$ とし, 1 のとき学生 s は研究室 l に配属されたことを表す. ある研究室 l に配属された学生の数を $AssignCount(l)$ で表し, $AssignCount(l) = \sum_{s \in S} Assign(s, l)$ とする.

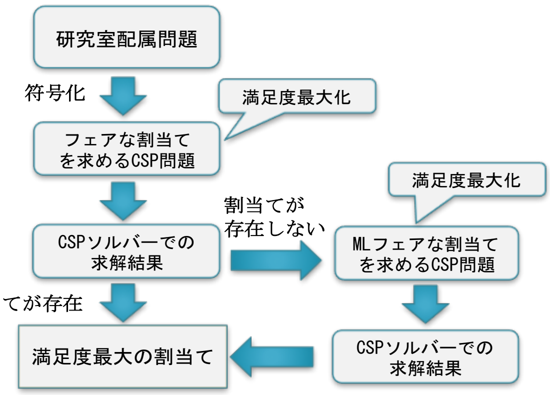

図 1: CSP 符号化手法の手順

4.1 無駄がない

ある学生 $s \in S$ の研究室 l_1, $l_2 \in L$ に対する選好順位を $l_1 \succ l_2$ としたとき, 学生 s の配属について"無駄がない"を表す制約は以下のようになる. 行きたい研究室が上限に達している状態もしくは, 今いる研究室から移動することで研究室の下限を下回るとき, 学生 s はより選好の高い研究室には移れないことを表す.

$$\forall s \in S, \forall l_1, l_2 \in L\ (l_1 \succ_s l_2) \wedge (Assign(s, l_2) = 1) \rightarrow$$
$$(AssignCount(l_1) = \mathrm{ub}_{l_1} \vee AssignCount(l_2) = \mathrm{lb}_{l_2})$$

4.2 フェア

ある学生 s , $s' \in S$ の研究室 l_1, $l_2 \in L$ に対する選好順位を $l_1 \succ l_2$ としたとき, 研究室 l_1, l_2 での選好順位は $s \succ_{l_1} s'$, $s \succ_{l_2} s'$ のとき"フェア"を表す制約は以下のようになる.

$$\forall s, s' \in S, \forall l_1, l_2 \in L\ (l_1 \succ_s l_2) \wedge (s \succ_{l_1} s')$$
$$\wedge (Assign(s, l_2) = 1) \rightarrow (Assign(s', l_1) = 0)$$

4.3 ML フェア

ML フェアの符号化はフェアの符号化と同様の制約で表す．4.2 節における学生 s, s' について，ML において $s \succ_{ML} s'$ であるとき，”ML フェア”を表す制約は以下のようになる．

$$\forall s, s' \in S, \forall l_1, l_2 \in L\, (l_1 \succ_s l_2) \wedge (s \succ_{l_1} s')$$
$$\wedge (Assign(s, l_2) = 1) \rightarrow (Assign(s', l_1) = 0)$$

4.4 満足度

フェアまたは ML フェアな解は無数に存在する場合があるため，それらの解の中で各学生がより満足する解を選択する必要があると考える．それらの解を選ぶ基準として満足度を定義する．

各学生は配属する研究室の評価によって満足度を決定するものとし，本論文では各学生の満足度を以下のように定義する．研究室の数を n とするとき，最も選好順位の高い研究室に割り当てられた場合，満足度を n とし，2 番目であれば *n-1* とする．最も選好順位が低い研究室満足度は 1 である．また割当ての解が複数ある場合には解の中で学生全体の満足度が最大となるものを解とする．そのため割当て M における学生 s の満足度を求める関数 $r(s, M)$ を用意し，目的関数を学生の満足度の総和 $\sum_{s \in S} r(s, M)$ として，これを最大化する．

満足度という尺度に基づき割当てを評価することで耐戦略性が失われてしまう．しかし，学生が嘘の選好順位を示して希望する研究室へ配属されるためには，その研究室に関わる学生の希望や研究室内での他の学生の順位 (成績) などをすべて把握しなければ有利になることは難しい．そのため本論文では，フェアかつ無駄がない解集合の中から解を 1 つ選出するための基準として，学生の満足度の総和を利用する．

5. 同順位を含む研究室配属問題

学生が持つ選好順位に同順位を認めることで，今までの全研究室に順序付けを行っていた場合には存在しなかった割り当ての望ましさが現れる．

5.1 同順位空きシート

ある割当て M において学生 $s' \in S$ が研究室 $l' \in L$ に配属され，l' は上限に達し，さらに s' の選好で l' と同順位の研究室 $l'' \in L$ $(l' =_{s'} l'')$ に空きが存在したとする．l よりも l' に行きたい学生 $s \in S$ は，l' における選好順位で s' よりも低い場合，配属することはできない．しかし図 2 のように，s' が同順位である l' から l'' へ移動することで空きが生まれ，s を l' に配属することができる．このように，ある学生に現在の研究室と同順位の研究室に移動してもらうことによって生まれる空きを要求することを**同順位空きシートを要求する**と定義する．形式的には下記の条件を満たす場合，学生 s は学生 s' に同順位空きシートを要求する．

$$\exists\, l, l' \in L, \exists\, s' \in S \begin{pmatrix} l' \succ_s l \\ \mid M(l') \mid = \mathrm{ub}_{l'} \\ s' \in M(l') \\ l' =_{s'} l'' \\ \mid M(l'') \mid < \mathrm{ub}_{l''} \\ \mid M(l) \mid > \mathrm{lb}_l \end{pmatrix}$$

同順位空きシート要求をする学生が存在しない割り当てを，**同順位に関する無駄がない**という．また，s' が l' と空きのある研究室 l'' の選好を同順位 $(l' =_{s'} l'')$ としなくても，別の学生 s'' が l'' に同順位 $(l'' =_{s''} M(s''))$ を持ち，s' が $M(s'')$ に同順位 $(l' =_{s'} M(s''))$ を持つことで複数の学生をまたいで，同様に同順位空きシートを要求できる．

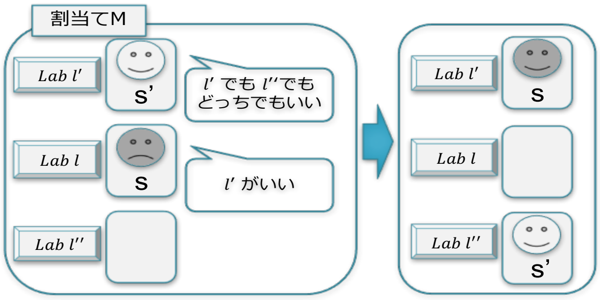

図 2: 同順位空きシートの例

5.2 CSP 符号化

学生 $s, s' \in S$ と，研究室 $l_1, l_2, l_3 \in L$ に対し，$l_1 \succ_s l_2 \succ_s l_3$，$l_1 =_{s'} l_3 \succ_{s'} l_2$ とする．さらに研究室 l_1, l_2, l_3 での選好順位は $s \succ_{l_1} s'$，$s \succ_{l_2} s'$，$s \succ_{l_3} s'$ のとき”同順位に関する無駄がない”を表す制約は以下のようになる．

$$\forall s, s' \in S, \forall l_1, l_2, l_3 \in L\, (l_1 \succ_s l_2 \succ_s l_3) \wedge (l_1 =_{s'} l_3 \succ_{s'} l_2) \wedge$$
$$(s \succ_{l_1} s') \wedge (s \succ_{l_2} s') \wedge (s \succ_{l_3} s') \wedge (Assign(s, l_2) = 1) \rightarrow$$
$$(\, (Assign(s', l_1) = 0) \vee (AssignCount(l_3) = \mathrm{ub}_{l_3}) \vee$$
$$(AssignCount(l_2) = \mathrm{lb}_{l_2}) \vee (AssignCount(l_1) \neq \mathrm{ub}_{l_1})\,)$$

5.3 同順位空きシートと満足度

上記で同順位に関する無駄がないを表す制約について述べたが，今回は学生全体の満足度を最大化する制約を追加することで，同順位に関する無駄がないの制約については必要がなくなることを証明する．そのため今回は同順位に関する無駄がないの制約は実験では追加せず行った．下記の命題を証明する．

命題． 学生全体の満足度を最大にする場合，同順位空きシート制約は必ず充足する．

証明． いま割当て M がフェア (または ML フェア) かつ無駄がなく学生全体の満足度が最大化されていると仮定する．このとき同順位空きシートを要求する学生がいると仮定すると，同順位空きシートを要求することから，ある学生 $s \in S$ は今いる研究室よりも選好の高い研究室に同順位空きシートで配属可能である．配属することでその学生 s の満足度は増加し，学生全体の満足度も同時に増加する．しかし，これは割当て M が満足度最大化されていることと矛盾する．よって，学生全体の満足度を最大化するとき，同順位空きシートを要求する学生が存在する割当ては存在しない．　証明終わり

6. 評価実験

全順序の問題と，同順位を含んだ問題の CSP 符号化手法について比較実験を行う． 今回は求解時間と，学生の割り当て結果について比較する． 問題は学生数を 50 から 550 までの範囲で 50 刻みに変化させたもので，研究室数は 10，各研究室の上限は学生数/研究室数+2 までとし，下限は 1 で固定した．また，学生の研究室に対する選好順位には相関があると考え，文献 [5] と同様に相関を設定した． この条件で各学生数の問

題を 50 問を生成して実験を行った．実験環境は, CPU Intel Xeon Processor E3-1230 V2(3.30GHz) , メモリ制限を 6GB, 1 問あたりの制限時間を 3600 秒で行う．また CSP ソルバーとして Sugar [6] を使用し, Sugar で使う SAT ソルバーとして glueminitsat[7] を使用した．(各バージョン Sugar は v2-2-1, glueminisat は 2.2.8)

図 3 は, 全順序と第 3 希望まで記入した場合及び MSDA の平均求解時間を示す．全順序は 10 個の研究室すべてを順位づけした場合を意味し, 第 3 希望まで記入は, 上位 3 つの研究室を順位づけし, 4 位以降を同順位とした場合を意味する．また MSDA は全順序と同様とする．全順序についてグラフが 300 人以降で切れているのは, 300 人以降の問題が 3600 秒以内に解けなかったことを表す．

図 3 の結果より, MSDA は非常に高速であることがわかる．同順位を含む問題の求解速度は MSDA には及ばないが, 全順序に比べ明らかに高速化し, さらに全順序が 300 人までに対し 550 人の問題まで求解可能となった．これは同順位を含むことで, 妥当な不満を表す制約が減少したこと, 満足度を表す変数ドメインが同順位により縮小したことが要因と考えられる．

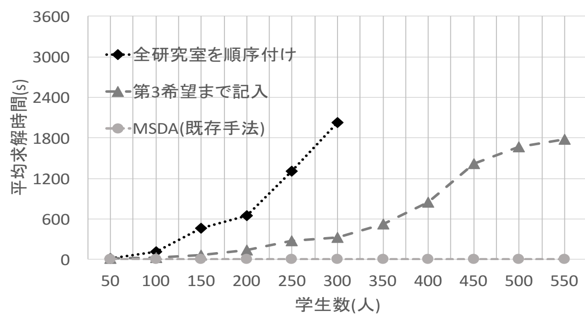

図 3: 平均求解時間の比較

図 4 は各学生が配属された研究室の希望順位の累計を示し, グラフの読み方は例えば横軸が 2, 縦軸が 50 のとき, 第 1 希望もしくは第 2 希望に配属された学生が 50%いることを表す．第 1 希望まで記入とは第 1 希望を決め, 2 位以降は同順位とした場合を意味する．同様に第 2 希望まで記入, 第 3 希望まで記入等も希望数まで順位づけをし, それ以降を同順位として設定した結果を表している．

学生の配属先の比率について, CSP 符号化手法は MSDA と比べると, 満足度を最大化することによって高い希望順位の研究室に配属できる学生の割合が増加したことが確認できる．さらに全順序の場合に比べ, 同順位を含む場合は, より選好順位の高い研究室に行ける学生の割合が増加している．これは, 同順位を認めることで妥当な不満を持つ場合が減少したことが考えられる．具体的には, 全順序の場合には多くの学生がより希望順位の高い研究室に配属できたとしても, 成績の低い下位の希望において, 妥当な不満を持つ学生が現れることで制約違反となり, 割当てとして出力されることがなかった．しかし, 第 n 希望まで順位づけをして, $n+1$ 位以降は同順位にすることでそういった下位で起こる妥当な不満が減少したことでより良い割当てが発見できたと考えられる．また同順位空きシートを要求できるようになったことももう一つの要因と考えられる．

7. まとめと今後の課題

本論文では研究室配属問題に対する CSP 符号化手法を示した．その特徴は, (1) フェアな解に対する完全性, (2) 満足度の

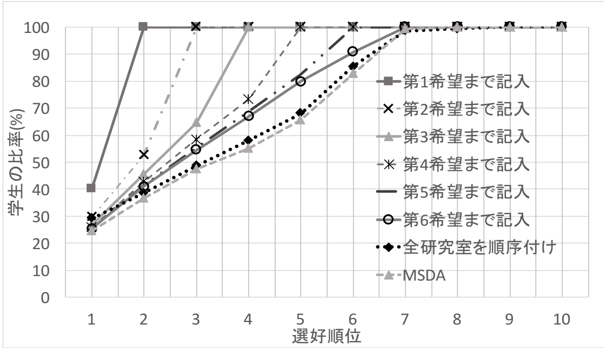

図 4: 学生の配属先選好順位の比率　学生 300 人

導入による解の評価, (3) 学生の研究室に対する選好順位に同順位を認めることにある．評価実験の結果は, 同順位の導入によって求解の高速化と, より上位に希望する研究室に配属される学生の増加が達成することができることを示している．

今後の課題として, 研究室が持つ選好順位にも同順位を認めた場合や, より多人数の問題のためのさらなるスケーラビリティの向上が挙げられる．また, 今回学生を中心としてその満足度を最大化することを考えたが, 研究室側についても同様の満足度について考慮するべきだと考えられる．

参考文献

[1] S.Ueda, D.Fragiadakis, A.Iwasaki, P.Troyan, and M.Yokoo.: Strategyproof Matching with Minimum Quotas, 2012. mimeo (an extended abstract version appeared in AAMAS, pages 1327-1328, 2012)

[2] Ehlers, L., Hafalir, I. E., Yenmez, M. B. and Yildirim, M. A.: School choice with controlled choice constraints Hard bounds versus soft bounds, Journal of Economic Theory, Vol.153, pp. 648-683, (2014).

[3] Gale, D. and Shapley, L. S.: College admissions and the stability of marriage, The American Mathematical Monthly, Vol.69, pp. 9-15, (1962)

[4] 田村 直之, 丹生 智也, 番原 睦則：制約最適化問題と SAT 符号化, 人工知能学会論文誌 25 巻 1 号 a(2010 年)

[5] Masahiro Goto, Naoyuki Hashimoto, Atsushi Iwasaki, Yujiro Kawasaki, Suguru Ueda, Yosuke Yasuda, and Makoto Yokoo.: Strategy-proof Matching with Regional Minimum Quotas, To appear in Proceedings of the Thirteenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014). May 2014.

[6] Tamura, N. and Banbara, M.: Sugar: A CSP to SAT Translator Based on Order Encoding, In Proceedingsof the 2nd International CSP Solver Competition, pp. 65-69,(2008).

[7] 鍋島 英知, 岩沼 宏治, 井上 克巳: GlueMiniSat 2.2.5: 単位伝搬を促す学習節の積極的獲得戦略に基づく高速 SAT ソルバー, コンピュータ ソフトウェア, Vol.29, No.4, pp. 4 146-4 160, (2012).