3A2-2

スケールフリーネットワーク上での非同期バックトラッキングの評価

Effect of DisCSP Variable-ordering Heuristics in Scale-free Networks

沖本 天太　岩崎 敦　横尾 真
Tenda Okimoto　Atsushi Iwasaki　Makoto Yokoo

九州大学大学院システム情報科学府
Graduate School/Faculty of Information Science and Electrical Engineering (ISEE), Kyushu University

A Distributed Constraint Satisfaction Problem (DisCSP) is a constraint satisfaction problem in which variables and constraints are distributed among multiple agents. Various algorithms for solving DisCSPs have been developed, which can be applied to any network structure. However, there exists virtually no work on examining the effect of particular network structures in DisCSPs.

In this paper, as an initial step toward developing specialized algorithms for particular network structures, we examine variable-ordering heuristics in scale-free networks. We use the classic asynchronous backtracking algorithm as a baseline algorithm and examine the effect of variable-ordering heuristics. First, we show that the choice of variable-ordering heuristics is more influential in scale-free networks than in random networks. Furthermore, we develop a novel variable-ordering heuristic that is specialized to scale-free networks. Experimental results illustrate that our new variable-ordering heuristic is effective and can reduce the required cycles by 30% compared to a standard degree-based variable-ordering heuristic.

1. はじめに

人工知能で扱われる問題の多くは，探索という要素を含んでおり，探索／制約充足問題として定式化できる．制約充足問題とは，有限で離散的な領域から値をとる複数の変数に制約を満たす値を割り当てる問題である．分散制約充足問題とは，制約充足問題における変数および制約が，複数のエージェントに分散された問題である [Yokoo 00]．分散制約充足問題では，各エージェントは自分の持つ変数の値を決定しようとするが，異なるエージェントの持つ変数間に制約があり，すべてのエージェントの変数への値の割当てがすべての制約を満たす必要がある．また，分散制約充足問題は変数をノード，制約をノード間のリンクに対応させることでネットワーク（制約ネットワーク）を用いて表現できる．

近年スケールフリーといわれるネットワーク構造が発見され [Barabási 03]，インターネット，タンパク質相互作用，SNS，論文引用ネットワークなど現実世界の様々なネットワークはスケールフリー構造であることがわかってきた．スケールフリーネットワークとは，一部のノードが膨大なリンクをもち，ほとんどのノードはごくわずかなノードとしかリンクしていないようなネットワークを指す．

制約充足問題では，このようなネットワーク構造に着目した研究がいくつか存在する．例えば，制約充足問題の応用問題として，レジスタ割当て問題，time-tabling，quasi-group 問題がネットワーク構造としてスモールワールド性をもつことに着目した研究 [Walsh 01] や，また周波数帯域割り当て問題や大規模な論理回路の検証で，ネットワーク構造にスケールフリー性が存在するものがあることに着目した研究がある [Devlin 09]．

一方，分散制約充足問題では，このようなネットワーク構造に着目した研究は著者らの知る限りほとんどない．制約充足問題に関するこのような応用問題で変数や制約が複数のエージェントに分散された問題を考えると，分散制約充足問題が適用できる．さらに，SNS で個人の代理人としてエージェントが存在する状況を考える．SNS はスケールフリー構造をもつ傾向があり，そのようなネットワーク構造でミーティングスケジュール問題を解決するとき，分散制約充足問題が適用できる．

我々は特定のネットワーク構造に特化したアルゴリズム／ヒューリスティックの開発を目的とする．本論文はその第一歩として，スケールフリーネットワークにおける変数順序付けヒューリスティックを提案し，その有効性を示す．本論文では分散制約充足問題の解法として古典的な非同期バックトラッキングを使用する．このアルゴリズムは，分散制約充足問題を解く最初の非同期かつ完全な解法として提案された [Yokoo 00]．分散制約充足アルゴリズムとして様々な効率的かつ洗練されたアルゴリズム（APO [Mailler 06], ABT-DO [Zivan 06] など）があるが，本論文ではスケールフリーネットワークに特化した変数順序付けヒューリスティックを開発するため，簡単な非同期バックトラッキングを用いることとする．

はじめに，スケールフリーネットワーク構造をもつ分散制約充足問題における非同期バックトラッキングの性能をシミュレーションで評価する．具体的には，エージェントの優先順位を変えながら，非同期バックトラッキングの性能をスケールフリーネットワークとランダムネットワーク上で評価する．

次にスケールフリーネットワークに特化した優先順位を決定するヒューリスティックを提案し，優先順位が次数順に基づくヒューリスティックと比較し，その有効性を示す．

連絡先: 沖本天太，九州大学大学院システム情報科学府，812-0395 福岡県福岡市西区元岡 744 番地，(092)802-3576，tenda@agent.is.kyushu-u.ac.jp

2. 分散制約充足問題の定式化

制約充足問題（CSP）は一般に次のように定義される．m 個の変数 $x_1, x_2, ..., x_m$ と，それぞれの変数が値をとる有限で離散的な領域 $D_1, D_2, ..., D_m$，および制約の集合が存在する．制約は述語論理によって内包的に定義されるとする．すなわち，制約 $p_k(x_{k1}, ..., x_{kj})$ は，直積 $D_{k1} \times ... \times D_{kj}$ に対して定義され，これらの変数の値が互いに整合のとれている場合に真となる．制約充足問題の解を求めることは，すべての制約を満

足する変数の値の組を求めることである．

分散制約充足問題（DisCSP） [Yokoo 00] とは，制約充足問題における変数および制約が，複数のエージェントに分散された問題である．各エージェントは自分の持つ変数の値を決定しようとするが，異なるエージェントの持つ変数間に制約があり，すべてのエージェントの変数への値の割当てがすべての制約を満たす必要がある．形式的には，エージェントの集合 {1,2,...,m} が存在し，各変数 x_j に対して，その属するエージェント i が定義される (belongs(x_j,i) と書く)．制約に関する情報も同様にエージェント間に分散される．エージェント i が制約 p_k を知っていることを know(p_k,i) と書く．

次の場合に，分散制約充足問題が解けたという．

- すべてのエージェント i において，$\forall x_j$ belongs(x_j,i) について，x_j の値が d_j に決定される．そして，すべてのエージェント k について，$\forall p_l$ know(p_l,k) なる制約が，x_1=d_1, x_2=d_2,....., x_n=d_n のもとで真となる．

制約充足問題の例としてよく用いられるものに n クイーン問題がある．各クイーンに対応する独立なエージェントが存在し，それらのエージェントが自分のクイーンの位置を決定しようとしているとする．このとき，この問題は分散制約充足問題として定式化される．

2.1 非同期バックトラッキング

非同期バックトラッキング（ABT）は分散制約充足問題を解く完全な解法として提案された [Yokoo 00]．このアルゴリズムでは, エージェント間の優先順位が前処理として決定される．エージェントは非同期, 並行に, 各自の局所的な情報に基づいて動作し, 新しい制約条件を互いに通信しあうことにより制約を満足する値の組み合わせを得る．各エージェントは並行して自分の持つ変数の値を一時的に決定し, その値を関連するエージェントに送信する．受信した値と制約条件違反が生じる場合には, エージェント間の優先順位を用いて, 優先順位の低いエージェントから値の変更が行われる．優先順位の低いエージェントにおいて, 優先順位の高いエージェントの変数の値と制約を満たす値が存在しない場合には, 逐次的なバックトラッキングが行われるのではなく, 新しい制約条件を導出し, 優先順位の高いエージェントに送信することにより, 非同期, 並行的なバックトラッキングが実現される．

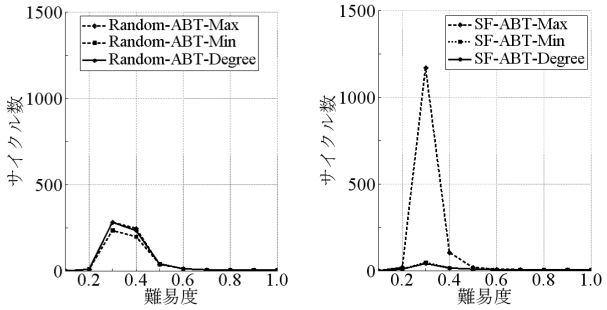
3. スケールフリーネットワーク

近年スケールフリー [Barabási 03] といわれるネットワーク構造が発見され，インターネット，タンパク質相互作用，SNS, 論文引用ネットワークなど現実世界の様々なネットワークがスケールフリー構造であることがわかってきた．従来これらのネットワークはランダムグラフとしてモデル化されていた．ランダムグラフとは，Erdös と Rényi の定義によれば，任意のノードの組が一定確率で隣接するグラフを指す（ER モデル）．このグラフでは，リンクがランダムにはられ，大半のノードはほぼ同数のリンクをもつ．

スケールフリーネットワークとは，ノードの次数分布がベキ乗則に従うネットワークを指す．次数 k をもつノードの出現頻度を $p(k)$ とし，ベキ乗則を以下の式で表す．γ はベキ指数とする．

$$p(k) \propto k^{-\gamma}$$

スケールフリーネットワークでは, 一部のノードが膨大なリンクをもち，ほとんどのノードはごくわずかなノードとしかリンクしていない．また, 他のノードに比べ極端に多くのリンクをもつノードのことをハブと呼ぶ．



(a) ランダムネットワーク（r_1）での非同期バックトラッキングの評価.

(b) スケールフリーネットワーク（sf_1）での非同期バックトラッキングの評価.

図 1: ランダムネットワークとスケールフリーネットワークでの非同期バックトラッキングの評価.

4. スケールフリーネットワークにおける変数順序付けの影響

本章では，エージェントの優先順位を変えながら，非同期バックトラッキングの性能をランダムネットワークとスケールフリーネットワーク上で評価する．

ここでは，離散イベントのシミュレータを用いて，エージェントの並行動作のシミュレートを行う．すなわち，各エージェントは自分自身のクロックを管理し，一つの制約チェックを行うたびにクロックの値を一つ増加させる (これを 1 サイクルと呼ぶ)．また，以後異なる難易度の問題を解く際に現れる局所的なサイクル数が増加するポイントをピークと呼ぶ．

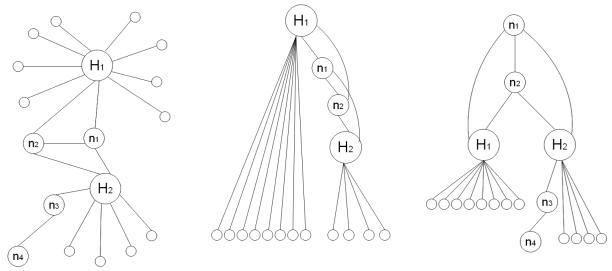
本論文では，スケールフリーネットワーク形成のツールとして Java プログラムを使用した[*1]．このプログラムは，ノード数，ベキ指数，最低次数を与えることでスケールフリーネットワークを生成することができる．ここで最低次数とは，各エージェントのもつ最低リンク数のことを指す．

実験では，10 個のランダムネットワーク（$r_1, \ldots, r_{10}$）と 10 個のスケールフリーネットワーク（$sf_1, \ldots, sf_{10}$）を生成した．各ネットワークで制約問題の難易度を 0.1 から 0.9 とし，各難易度で 100 個のランダムな制約問題を解かせ，サイクル数の平均値を求めた．また各ネットワークでエージェントの優先順位をランダムに 10 通り 決定した．スケールフリーネットワークは，ノード数＝ 100，最低次数＝ 2，ベキ指数 $\gamma = 1.8$ のパラメータで生成した [*2]．ランダムネットワークは，スケールフリーネットワークと同じノード数とリンク数になるように生成した．簡単のため，ここではドメイン数＝ 3 とした．

図 1 に，ランダムネットワーク（r_1）とスケールフリーネットワーク（sf_1）における実験結果を示す．ここでは，3 つの特徴的な結果を示す優先順位を与えた場合を選んでいる．Random-ABT-Max (Min) と SF-ABT-Max (Min) は，それぞれのネットワークにおけるピーク時のサイクル数が最大（最小）とな

∗1 Owen Densmore, “An Exploration of Power-Law Networks,” (http://backspaces.net/sun/PLaw/index.html)

∗2 γ= 1.8, 2.2, 2.6, 3.0 の場合でも本質的な結果は変わらなかった.



(a) ある分散制約充足問題の制約ネットワーク．H_1, H_2 をハブとする．

(b) 次数順による優先順位によって決定される疑似木．H_1 の次数が最も大きく，疑似木のルートに位置している．

(c) ALH による優先順位によって決定される疑似木．H_1, H_2 は異なるブランチに位置している．

図 2: 疑似木の例

る非同期バックトラッキングのサイクル数を表わしている．また，Random-ABT-Degree と SF-ABT-Degree は優先順位が次数順（大きい）に従う非同期バックトラッキングのサイクル数を表わしている．

ランダムネットワークでは優先順位が変わってもピーク時のサイクル数はほとんど変化しないのに対し，スケールフリーネットワークではピーク時のサイクル数に大きな差がでた．実際，図 1 では与えたネットワーク構造に関わらず，難易度が上がるにつれサイクル数はいったん増加し，難易度 0.3 付近でピークをむかえ減少しているのがわかる．このことは難易度 0.3 までの問題には解が存在し，難易度 0.3 の問題が解の存在する最も難しい問題であることを示している．また難易度 0.4 以上では解なしという答えを出している．図 1(a) より r_1 では，ピーク時のサイクル数は 235 から 283 まで変化した．一方，図 1(b) より sf_1 では，ピーク時のサイクル数は 47 から 1171 まで大きく変化した．

以上の実験結果から，スケールフリーネットワークでは，エージェントの優先順位がアルゴリズムの性能に大きな影響を与えることがわかった．これらの結果は他のネットワーク（$r_2, \ldots, r_{10}, sf_2, \ldots, sf_{10}$）でも変わらなかった．また，さらにピーク時のサイクル数がもっとも小さくなるとき，その優先順位はノードの次数順に近くなっていた．

5. スケールフリーネットワークに特化した変数順序付けヒューリスティックの提案

本章ではスケールフリーネットワーク構造に特化した変数順序付けヒューリスティック ALH（variable ordering with Average Length between Hubs）を提案する．

5.1 ALH の概要

$G = (N, E)$ をベキ乗則グラフとする．$N = \{n_i | i \in \mathbb{N}\}$ をノードの有限集合，$E = \{e(n_i, n_j) | n_i, n_j \in N, n_i \neq n_j\}$ をエッジの有限集合とする．非同期バックトラッキングでは，解を得るために事前に変数の優先順位が決定されている必要がある．変数の優先順位が決定されると，優先順位を基に疑似木が作成される．ALH によって決定された優先順位では，ハブは疑似木内で異なるブランチに位置する．

ある定数 $c \in \mathbb{N}$ 以上のエッジをもつノードのことをハブと呼ぶ．H をハブの集合とし以下のように定義する．

$$H = \{n_i | n_i \in N, deg(n_i) \geq c\}.$$

$deg(n_i)$ はノード n_i の次数を表す．各エージェントは自分が H に属するかを知っているとする．

グラフ G 上の任意の 2 つのノード間の最短距離を $dis(n_i, n_j)$ で表わし，H 内のハブまでの最短距離の平均を以下のように定義する．

$$\forall n_i \in N, n_i^{av} = \Sigma_{n_j \in H} dis(n_i, n_j)/|H|.$$

またハブ間の最短距離の平均値を以下のように定義する．

$$\forall n_i \in H, h^{av} = \Sigma_{n_i \in H} n_i^{av}/|H|.$$

条件式 $n_i^{av} \leq h^{av}$ を満たすノードの集合を境界集合 BS と呼び，以下のように定義する．

$$BS = \{n_i | n_i^{av} \leq h^{av}\}.$$

ここでエージェントの優先順位をどのように決定していくかをみる．各エージェントの優先度は境界集合を基に決定され，境界集合内のハブ以外のエージェントから高い優先度を付けていく．その際，最短距離の平均値が小さいものほど高い優先度を付ける．もし平均値が同じ場合は，両者の次数を比較し，次数の大きい方に高い優先度を付ける．次にハブに優先度を付ける．その他のエージェントに関しては，境界集合の要素までの最短距離の和によってその優先度を付けていく．つまり，ノード n_i に対して $\sum_{n_j \in BS} dis(n_i, n_j)$ を求め，その値の小さいものほど大きな優先度を付ける．ここでも同じ値をもつエージェントが存在する場合は，次数を比較し，次数の大きい方に高い優先度を付ける．

ハブ同士が直リンクでつながっていないスケールフリーネットワークについて考える．ALH によって決定されるエージェントの優先順位では，境界集合に属するエージェントが最も高い優先度をもつ．境界集合に属するハブ以外のエージェントは，ハブよりも高い優先度をもつ．このため ALH によって決定されたエージェントの優先順位を基に疑似木を作成すると，境界集合の要素が疑似木の上位に位置し，やや上位にハブが位置する．ハブはこの疑似木内では異なるブランチに位置しており，このような疑似木では，境界集合とハブをルートとした部分木のクラスターを考えることができる．このようなクラスターでは部分問題を別々に解くことができ，効率的と考えられる．ALH は，このようなクラスターをもつ疑似木を想定し，各エージェントの優先順位を決定するヒューリスティックである．一方，ハブ同士が直リンクでつながっているスケールフリーネットワークでは，境界集合の要素はハブのみとなり，各エージェントの優先順位はハブまでの最短距離の平均値によって決定される．このように決定されたエージェントの優先順位を基に疑似木を作成すると，優先順位が次数順に従ったものと類似する．また，ＡＬＨの計算量は各ハブまでの最短距離を求めさえすればよいので O(n) となる．よって、計算量はエージェントの優先順位をランダムに決定する従来の非同期バックトラッキングとほとんど変わらない．

ここで簡単な例を与える．図 2(a) に，ある分散制約充足問題の制約ネットワークを示す．$H_1, H_2, n_1, n_2, n_3, n_4$ をエージェントとし H_1, H_2 をハブとする．各エージェントの次数が $deg(H_1) > deg(H_2) > deg(n_1) > deg(n_2) > deg(n_3) > deg(n_4)$ で与えられているとする．図 2(b) に次数順による優

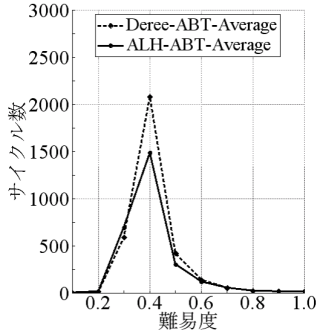


図 3: ALH と次数順による非同期バックトラッキングの評価.

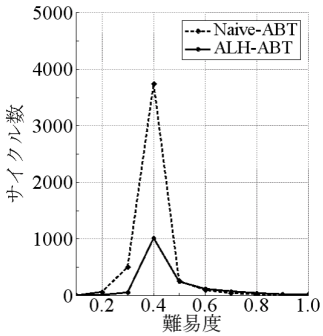


図 4: ALH と単純なヒューリスティックによる非同期バックトラッキングの評価.

先順位によって決定される疑似木を示す．この疑似木ではハブはルートか疑似木の上位に位置する．これに対し，ALH では以下の境界集合が定義される．

$$BS = \{H_1, H_2, n_1, n_2\}.$$

また各ノードの最短距離の平均値はそれぞれ $n_1^{av} = n_2^{av} = 1$, $n_3^{av} = 2$, $n_4^{av} = 3$, $h^{av} = 1$ と求まる．この値を基にエージェントの優先順位を決定すると $n_1, n_2, H_1, H_2, n_3, n_4$ のようになる（優先度の大きい順）． 図 2(c) に ALH によって決定された優先順位による疑似木を示す．この疑似木ではハブは異なるブランチに位置しているのがわかる．一般に，ALH による優先順位によって決定される疑似木では境界集合の要素が疑似木上位に位置し，ハブはやや上位に位置し，疑似木内で異なるブランチに位置する．

5.2 実験

実験では, スケールフリーネットワーク上で, ALH と次数順によって優先順位を決定した非同期バックトラッキングをシミュレーションで評価する．10 個のスケールフリーネットワークをノード数= 100，最低次数= 3，$\gamma = 1.8$ のパラメータで生成した．各ネットワークで制約問題の難易度を 0.1 から 0.9 とし，各難易度で 100 個のランダムな制約問題を解かせその平均値を求めた．ここではドメイン数= 10 とした．実験結果は，10 個のスケールフリーネットワークに対し, 100 個のインスタンスを解かせた合計 1000 回の平均値を表す．図 3 にその実験結果を示す．Degree-ABT-Average は優先順位が次数順に従う非同期バックトラッキングを表し，ALH-ABT-Average は優先順位が ALH によって決定された非同期バックトラッキングを表している．特徴的なサイクル数の差はピーク時に現れ ALH の有効性が観察された．図 3 より，難易度 0.4 付近でピークを迎え, ALH-ABT-Average のピーク時のサイクル数の平均は Degree-ABT-Average の 2082 サイクルに対し，約 30 %減の 1489 サイクルであった．

以上の結果から，優先順位を ALH により決定したものは，優先順位が次数順に従うものよりアルゴリズムの性能を向上させることがわかった．著者らは ALH の有効性がその他のパラメータでも変わらないことを確認した．

6. 考察

本章では, ALH と境界集合を単純に定義したヒューリスティックを比較する．前章では, ALH がエージェントの優先順位を決定する有効なヒューリスティックであることを示した．しかし, ALH の特に境界集合の定義はやや複雑に見えるため，より単純なヒューリスティックでも同様の結果が得られるのではないかという疑問が残る．そこで単純なヒューリスティックを以下のように定義する．各 $n_i \in N$ に対し, 自分から最も近い 2 つのハブ $h_{i,1st}, h_{i,2nd} \in H$ までの最短距離を $dis_{1st}(n_i)$, $dis_{2nd}(n_i)$ で表す．以下の条件式を満たすノードの集合を単純な境界集合とする．

$$\text{単純な境界集合} = \{n_i \mid |dis_{1st}(n_i) - dis_{2nd}(n_i)| \leq 1\}.$$

単純な境界集合は, 自分から最も近い 2 つのハブのちょうど境界に位置するエージェントを含む．エージェントの優先順位は ALH 同様，まずは境界集合の要素から高い優先度を付けていく．境界集合に属さないその他のエージェントに関しては次数の大きいもの順に優先度を付ける．図 4 に単純なヒューリスティックと ALH を比較した実験結果を示す．スケールフリーネットワークはノード数= 100，最低次数= 3，$\gamma = 1.8$ のパラメータで生成した．Naive-ABT は単純なヒューリスティックによる非同期バックトラッキングのサイクル数を表している．Naive-ABT のピーク時の 3744 サイクルは，ALH-ABT の 1067 サイクルの約 3.7 倍大きかった．この結果から，ALH はスケールフリーネットワークに特化した有効なヒューリスティックであることがわかった．

7. おわりに

本論文では，非同期バックトラッキングにおけるエージェントの優先順位が，スケールフリーネットワークではランダムネットワークよりアルゴリズムの性能に大きな影響を与えることを示した．次に, スケールフリーネットワークに特化した変数順序付けヒューリスティック（ALH）を提案した．実験では優先順位が ALH と次数順に従う非同期バックトラッキングの性能を比較し，ALH の有効性を示した．

著者らは今後の研究課題として，動的なヒューリスティック/アルゴリズムの開発を考えている．

参考文献

[Barabási 03] Barabási, A.-L.: Linked: The New Science of Networks, *J. Artificial Societies and Social Simulation*, Vol. 6, No. 2 (2003)

[Devlin 09] Devlin, D. and O'Sullivan, B.: Preferential Attachment in Constraint Networks, in *ICTAI*, pp. 708–715 (2009)

[Mailler 06] Mailler, R. and Lesser, V. R.: Asynchronous partial overlay: A new algorithm for solving distributed constraint satisfaction problems, *Journal of Artificial Intelligence Research (JAIR)*, Vol. 2005, p. 2006 (2006)

[Walsh 01] Walsh, T.: Search on High Degree Graphs, in *IJCAI*, pp. 266–274 (2001)

[Yokoo 00] Yokoo, M. and Hirayama, K.: Algorithms for Distributed Constraint Satisfaction: A Review, *Journal of Autonomous Agents and Multi-agent Systems*, Vol. 3, No. 2, pp. 189–211 (2000)

[Zivan 06] Zivan, R. and Meisels, A.: Dynamic Ordering for Asynchronous Backtracking on DisCSPs, *Constraints*, Vol. 11, No. 2-3, pp. 179–197 (2006)