4D1-4

サンプリングに基づく分散制約最適化手法の複数コンテキスト化と探査戦略の検討

A Study on Multiple Contexts and Search Strategies for Sampling Based DCOP Solver

松井 俊浩　松尾 啓志
Toshihiro Matsui　Hiroshi Matsuo

名古屋工業大学
Nagoya Institute of Technology

Distributed Constraint Optimization Problem is a fundamental problem in multiagent cooperation. To solve large scale and complex problems, several inexact solution methods have been proposed. In this study, we focus on an inexact solution method based on Gibbs sampling. Since this solution method is a synchronous distributed algorithm, a single sampling process is iteratively performed by an entire group of agents. Moreover, relatively long rise time of the objective value is necessary due to a probabilistic sampling. We propose several efficient methods, where agents process multiple contexts of sampling and cooperate for a local search strategy.

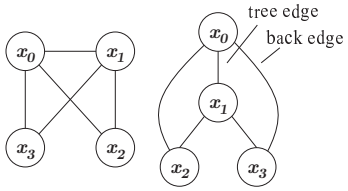
1. はじめに

分散制約最適化問題 (DCOP) [Modi 05, Petcu 05, Mailler 04, Farinelli 08, Rogers 11, Okimoto 11, Vinyals 11, Zivan 08, Ottens 12, Nguyen] はマルチエージェントシステムにおける基本的な最適化問題として研究されている．分散センサ網や電力資源の資源割り当て，会議スケジューリング，災害応答を含む応用的な問題が DCOP として定式化される [Zhang 05, Miller 12, Maheswaran 04, Ramchurn 10]．DCOP の解法は厳密解法 [Modi 05, Petcu 05, Mailler 04] と非厳密解法 [Farinelli 08, Rogers 11, Okimoto 11, Vinyals 11, Zivan 08, Ottens 12, Nguyen] に大別される．厳密解法は最適解を得るが，変数や制約の数および制約の密度に関するパラメータについて，時間，空間，メッセージに関する複雑度のいずれかが指数関数的であり，大規模で複雑な問題における処理コストの増加が無視できない．そのため，大規模かつ複雑な問題の解法としては，処理コストが比較的小さい非厳密解法が用いられる．非厳密解法には，問題を近似する手法 [Rogers 11, Okimoto 11]，局所的な最適化に基づく手法 [Vinyals 11, Zivan 08] および，確率的探索に基づく手法 [Zivan 08, Ottens 12, Nguyen] などがある．これらのうち，確率的なサンプリングに基づく探索手法である分散 Gibbs サンプリング [Nguyen] は，空間複雑度が問題の規模について線形であり，理論的には最適解への収束が保証されている．また，独特の分散スナップショットアルゴリズムと統合されているため，実時間性が要求される場合に有用である．しかし，従来手法は同期型の分散アルゴリズムであり，単一のサンプルを取得するたびに系全体を駆動するため，通信のオーバヘッドが大きい．また，確率的サンプリングには冗長さがあると考えられる．そこで，本研究では複数のサンプリングを並行する手法とその探索の戦略について検討する．

2. 準備

2.1 分散制約最適化問題

分散制約最適化問題 (DCOP) は，エージェントの集合 A，変数の集合 X，変数の値域の集合 D，制約の集合 C および各制約に対応する関数の集合 F により定義される．エージェント $i \in A$ は自身の状態や意思決定を表す変数の集合 $X_i \subseteq X$ を持つ. 変数 $x_m^i \in X_i$ は離散値の集合 $D_m^i \in D$ に含まれる値をとる. 変数 x_m^i の値はエージェント i によってのみ決定される. 制約 $c_j \in C$ は変数間の関係を表す. 制約 c_j により関係する各変数値の割り当て $\mathcal{A}_j = \{(x_0^j, d_0^j), \cdots, (x_k^j, d_k^j)\}$ の評価値は，関数 $f_j(\mathcal{A}_j) : D_0^j \times \cdots \times D_k^j \rightarrow \mathbb{R}_+$ により定義される．目的は，大域的な評価値 $\sum_{f_j \in F,\ \mathcal{A}_j \subseteq \mathcal{A}} f_j(\mathcal{A}_j)$ を最大化する最適解 $\mathcal{A}^*$ を求めることである．以降では，一般に基本的な問題とされる，各エージェント i が単一の変数 x_i を持ち，各制約が二項制約である場合を対象とする．また，表現を簡単にするために，変数 x_i とエージェント i を区別せずに扱う．各エージェントは，自身の変数に関する制約と関数を知る．最適解を求める計算では，メッセージ通信に基づく分散アルゴリズムを用いる．



(a) constraint network　(b) pseudo-tree

図 1: 疑似木

2.2 擬似木

変数と二項制約からなる問題は制約網により表現される．疑似木 [Petcu 05] は制約網の生成木に対応するグラフ上の構造である．典型的な疑似木は制約網に対する深さ優先探索木に基づく．たとえば図 1(a) の制約網から図 1(b) の疑似木が生成される．生成木に対応する辺を木辺と呼び，それ以外の辺を後退辺と呼ぶ．疑似木の各頂点を，生成木の頂点とみなし，根，葉，親，子などの用語により頂点間の関係を表す．疑似木の異なる部分木の間には辺が無いため，各部分木における計算を並行できる．

2.3 分散 Gibbs サンプリング

分散 Gibbs サンプリング [Nguyen] は，擬似木と Gibbs サンプリングに基づき確率的な探索を行う DCOP の解法である．この手法では図 2(a) に示される 2 種類のメッセージが用いられる．

連絡先: 松井 俊浩，名古屋工業大学，〒 466-8555 愛知県名古屋市昭和区御器所町，matsui.t@nitech.ac.jp

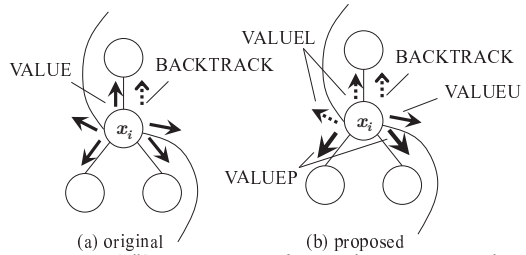


図 2: 分散 Gibbs サンプリングのメッセージ

初期状態では，各エージェントは制約で関連する近傍のエージェントの変数値を知る．各エージェントは擬似木に基づき同期して反復的に解の探索を行う．各反復では先ず，根ノードから葉ノードに向かってエージェントの制御が伝搬する．このとき，各エージェントは自変数値について 1 回のサンプリングを行う．サンプリングは，Gibbs サンプリングに基づいて行われる．サンプリングの結果，これまでの解が改善される場合に，その解を保存する．また，サンプリングされた変数値を，制約で関連する近傍のエージェントに，VALUE メッセージにより通知する．このとき，子ノードにはこれまでのサンプリングによる評価値の変化量の合計値を通知し，制御を渡す．評価値の変化量は，子ノードが解の改善の有無を判定するために用いられる．

葉ノードまで処理が伝搬した後，葉ノードから根ノードに向かって BACKTRACK メッセージを用いて，エージェントの制御が伝搬する．このとき，各エージェントを根とする部分木における，サンプリングによる評価値の変化量を集計し，これまでの解が改善される場合に，その解を保存する．各エージェントは局所的に最適化を行うが，擬似木に基づく同期により排他制御されるため，評価値の差分のみを考慮して解の改善の有無を判定できる．このような反復により，確率的に解を探索する．この解法は最良解を保存するスナップショットアルゴリズムと統合されているため，根ノードに制御が戻った任意の時点で，木の深さのメッセージ伝達回数で，トップダウンに解を決定して終了できる．解法の詳細は [Nguyen] を参照されたい．

従来手法は同期型の分散アルゴリズムであり，単一のサンプルを取得するたびに，根ノードから葉ノードまでメッセージ伝搬を往復する．そのため，通信のオーバヘッドが大きいといえる．また，確率的サンプリングを局所的な探索として捉えると，ランダムな変数値の選択には冗長さがある．そのため，探索の初期段階の解品質には改善の余地があると考えられる．

3. 提案手法

3.1 複数のサンプリングの並行

同一のメッセージ伝搬において，複数のサンプリングを並行する．これらをサンプリング過程と呼ぶ．このために，それぞれのサンプリング過程に対応するデータ構造を複数用いる．また，メッセージも全てのサンプリング過程の情報を運ぶように拡張される．それぞれのサンプリング過程は，全てのエージェントに分割して共有され，メッセージ交換により情報が対応付けられる．複数のサンプリングを並行して大数の法則を活用することにより，解を改善する機会の増加が期待される．

このような方法では，時間と空間およびメッセージサイズのコストはサンプリング過程の数について線形に増加する．その一方で，メッセージ通信に要する時間が計算時間と比較して十分に大きい場合は，通信のオーバヘッドを隠蔽することが期待される．本研究ではこのような通信のオーバヘッドに関する条件を仮定し，同一の通信回数においてサンプリングを並行する場合の解の改善の程度に注目する．

3.2 評価値の高いサンプルの選択

複数のサンプリング過程を並行する際に，評価値が高い解を複製して優先的に探索することにより，解品質が向上する可能性がある．その一方で，サンプリング過程の総数が一定である場合は，このような解の選択は局所解への集中を招く可能性がある．このような戦略の効果を検討するために，ある回のあるサンプリング過程の結果を，次の回の別のサンプリング過程に引き継ぐ手法を導入する．

各サンプリング過程では，各エージェントは近傍の変数値，最良解，評価値の差分情報などを持つため，前回と今回のサンプリング過程の情報を選択し統合する必要がある．このために，トップダウンな制御の伝搬の際に各エージェントにおいて，疑似木の上位の部分については各サンプリング過程で今回に個別に更新された情報を用い，下位の部分については情報を引き継ぐ元のサンプリング過程の前回の情報を用いる．

評価値が高い解を複製する戦略として，次のような手法を用いる．先ず，複数のサンプリング過程の一部を，評価値が高い解を強調するためのエリートの集団とする．あるサンプリングの反復を始める際に，前回のサンプリングの際に評価値が高かった上位の結果をエリートの集団に引き継ぐ．このとき，引き継ぐ元がエリートではないサンプリング過程である場合は，そのサンプリング過程の結果は継続される．この方法では，エリートではないサンプリング過程において，比較的大きな解の改善があった場合に，そのサンプリング過程が複製されて強調される．

3.3 値域の制限

従来手法では Gibbs サンプリングに基づいて変数値を選択する．このような確率的な変数の選択には，解の改善の機会を無視するような，冗長な変数値の選択があると考えられる．そこで，サンプリングにおいてフィルタリングをすることにより，解の改善の機会を強制する手法を導入する．

変数値の選択を制御するために，特定の値域をマスクして制限する．各エージェントは変数値を選択する際に，マスクされていない変数値のいずれかを選択する．このような枠組とマスクの設定の戦略により，サンプリングに変化を与えることを試みる．

3.4 隣接エージェント間の協調

従来手法において，サンプリングよりも解の改良を優先することは，基本的には，局所的な問題について決定論的な山登り法を実行することである．このような方針は，「深い」局所解に収束する可能性を高めるため，解の確率的な改善の機会が小さくなる可能性があると考えられる．その一方で，探索の初期段階の解の改善を早める効果が期待される．

従来手法のサンプリングにおいては，各変数について，その変数と制約で関係する近傍の変数の，現在の値のみを用いて，変数値を選択する確率を計算する．このような方法では，相手のエージェントが変数値を変更する可能性についての知識を用いない．

そこで，擬似木に基づいて情報を交換し，隣接エージェント間の協調により解を改善する局所的探索を導入する．提案手法では，トップダウンな制御の伝搬とボトムアップな制御の伝搬のそれぞれに，協調のための処理を追加する．このために，変数値の更新を近傍のエージェントに同報する従来手法のメッセージを図 2(b) に示されるように変更する．提案手法では，これらのメッセージはトップダウンな段階で送信されるメッセージ (VALUEP, VALUEU)，およびボトムアップな段階で送信されるメッセージ (VALUEL, BACKTRACK) に分離される．さら

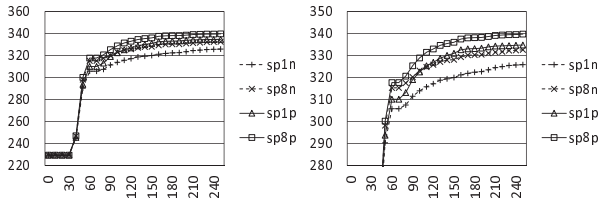


図 3: rnd, $c = 75$ (横軸: メッセージサイクル数, 縦軸: 評価値)

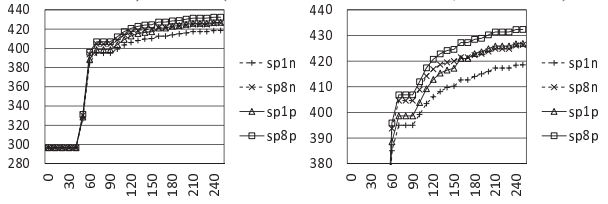


図 4: rnd, $c = 100$ (横軸: メッセージサイクル数, 縦軸: 評価値)

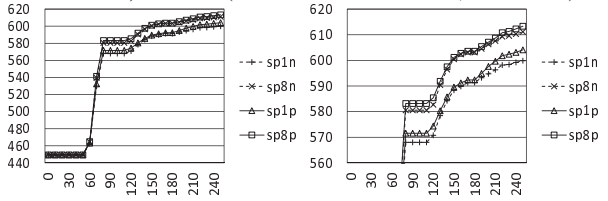


図 5: rnd, $c = 150$ (横軸: メッセージサイクル数, 縦軸: 評価値)

に，トップダウンな段階で送信されるメッセージは，親から子へのメッセージ VALUEP と，親から子以外の下位近傍のエージェントへのメッセージ VALUEU に分離される．

先ず，ボトムアップな段階では，各エージェントは上位の近傍の各エージェントに対して，VALUEL メッセージにより協調的な解の改善を提案する．このとき，現在注目している上位の近傍以外の近傍の変数値を固定した場合に，自身と相手のエージェントにより評価値が改善できる量を求める．この量は相手の変数値ごとに計算される．この改善量の表と，そのときに固定すべき他の近傍の変数の情報を，相手のエージェントに伝達する．

次にトップダウンな段階では，各エージェントは下位の近傍のエージェントから提案された改善量を考慮し，いずれかのエージェントと協調するか否かを判断する．このとき，下位のエージェントからの改善量は，従来の近傍の改善量と結合して評価される．組み合わせ探索を抑制するために，いずれかのエージェント一つと協調する．あるエージェントと協調することにより解が改善する場合，自身の変数値を解が改善する値のみ選択できるようにマスクする．また，そのエージェントと協調することを，VALUEP および VALUEU メッセージにより下位の近傍エージェントに伝達する，また各子エージェントには，協調のために固定すべき変数を VALUEP メッセージにより伝達する．

下位のエージェントは，自身が上位の近傍のエージェントが選択した協調の対象であれば，必ず提案した改善量となる変数値を選択する．協調の対象ではない他のエージェントは，上位から受信した固定すべき変数を考慮し，自身が下位のエージェントと協調する余地があれば，さらに協調を選択する．また，上位のエージェントの一部が協調を選択している状況では，他のエージェントは自変数値が固定されず協調もしない自由な状態であっても，現在の解と同等以上の評価値となる変数値のみ選択できるようにマスクする．

4. 評価

提案手法を実験により評価した．例題として，50 変数からなる次の 2 種類の DCOP を用いた．**rnd:** 関数値が一様分布に

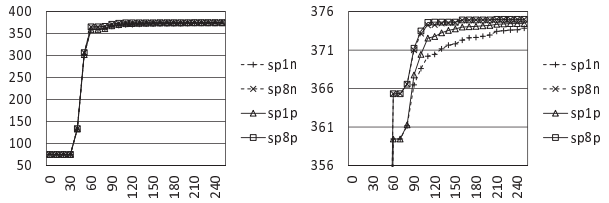


図 6: gcl, $c = 75$ (横軸: メッセージサイクル数, 縦軸: 評価値)

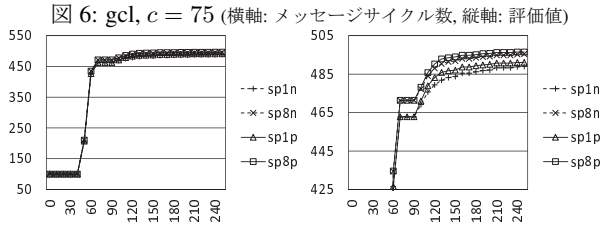


図 7: gcl, $c = 100$ (横軸: メッセージサイクル数, 縦軸: 評価値)

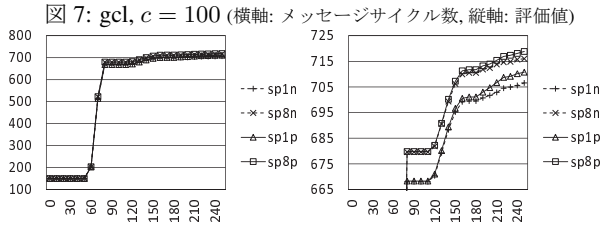


図 8: gcl, $c = 150$ (横軸: メッセージサイクル数, 縦軸: 評価値)

基づくランダムな $[1, 5]$ の整数値の問題．変数の値域のサイズは 5 とした．**gcl:** 制約違反と充足の関数値がそれぞれ 1 と 5 の 3 色の頂点彩色問題．制約の数 c は 75, 100 および 150 とした．

各パラメータについて 10 問の問題をそれぞれ 10 回試行した結果を評価した．複数回の試行では，サンプリングの戦略の影響に注目するために，各サンプリング過程の初期値は同一の値とした．

図 3 から 5 に，rnd の場合の結果を示す．各グラフの横軸は，同期されたメッセージ通信の回数である，メッセージサイクル数を表し，縦軸は系全体の評価値を表す．系全体の評価値として，分散 Gibbs が各回の反復を完了した時点における，大域的な評価値の下界値を用いた．

ここでは，次の手法を比較した．**sp1n:** 従来手法．**sp1p:** 隣接エージェント間の協調による解の改善を用いる手法．**sp8n, sp8p:** それぞれの手法を 8 個のサンプリング過程により並行する手法．

図 3 の制約数 $c = 75$ の場合には，提案手法 sp1p と sp8p の立ち上がりが，従来手法 sp1n および従来手法を並行する sp8n よりも早い傾向が見られた．特に，sp1p は比較的早い段階で sp8n を平均的に上回った．この原因は，比較的疎な問題では近傍エージェント間の協調による解の改善の機会が多いためであると考えられる．その一方で，c が比較的大きい場合の結果である図 4 および 5 では，サンプリング過程を並行することによる効果が比較的大きい傾向が見られた．

図 6 から 8 に，gcl の場合の結果を示す．これらの結果では，近傍エージェント間の協調による解の改善の効果は見られるものの，サンプリング過程を並行することによる効果が比較的大きい傾向が見られた．この原因は，従来手法のサンプリングによる解の改善が比較的有効であったためであると考えられる．

図 9 から 11 に，rnd の場合に，異なるサンプリング過程の組み合わせを比較した結果を示す．ここでは，次の手法を比較した．**sp8n, sp8p:** 従来手法と，隣接エージェント間の協調による解の改善を用いる手法それぞれを 8 個のサンプリング過程により並行する手法．**sp8epon** 隣接エージェント間の協調による解の改善を用いる手法 4 つをエリートとし，残りを従来手法とする手法．**sp8epop:** 全て隣接エージェント間の協調に

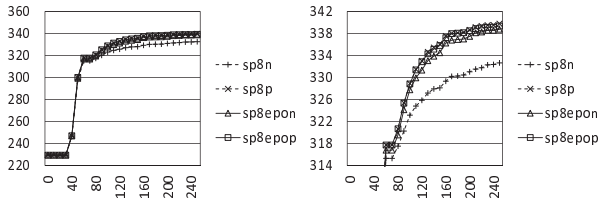


図 9: rnd, $c = 75$ (横軸: メッセージサイクル数, 縦軸: 評価値)

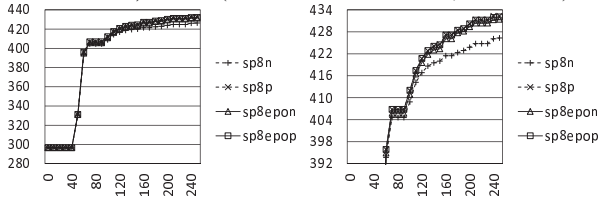


図 10: rnd, $c = 100$ (横軸: メッセージサイクル数, 縦軸: 評価値)

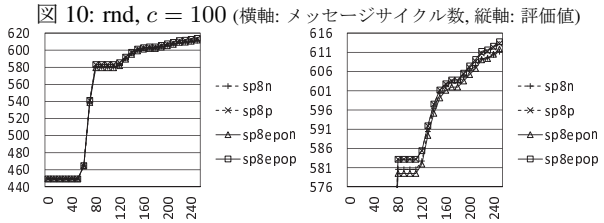


図 11: rnd, $c = 150$ (横軸: メッセージサイクル数, 縦軸: 評価値)

よる解の改善を用いる手法を用い，そのうちの 4 つをエリートとする手法．

これらの結果では，図 9 および 10 のように，$c = 75$ および 100 の場合に，sp8n よりも他の手法により得られる評価値が高い傾向が見られた．その一方で，他の手法の間に顕著な違いは見られなかった．これは，隣接エージェント間の協調による解の改善を行うサンプリング過程が含まれていることが，これらの結果において支配的であるためであると考えられる．また，sp8epon の評価値がやや低い傾向が見られた．この原因は，これらの問題における探索の初期段階では，従来手法の結果をエージェント間の協調を用いる手法により改善することよりも，エージェント間の協調を用いる手法に全てのサンプリング過程を割り当てることが有効であることを示唆している．sp8epop の評価値は sp8p よりも僅かに高い場合が見られた．これは，評価値が高い結果を複製して探索の機会を増やした影響の可能性があるが，より詳細な検討が必要である．

5. おわりに

本研究では，サンプリングに基づく分散制約最適化問題の改良手法を検討した．複数のサンプリングの過程を並行することにより，メッセージあたりの探索の効率を改善する手法および，エージェント間の協調的な解の改善を行う手法を導入した．複数のサンプリングの結果を協調的に活用する手法，擬似木の構造を活用する他の手法の導入，異なる確率的サンプリングの適用が今後の課題である．

参考文献

[Farinelli 08] Farinelli, A., Rogers, A., Petcu, A., and Jennings, N. R.: Decentralised coordination of low-power embedded devices using the max-sum algorithm, in *AAMAS08*, pp. 639–646 (2008)

[Maheswaran 04] Maheswaran, R. T., Tambe, M., Bowring, E., Pearce, J. P., and Varakantham, P.: Taking DCOP to the Real World: Efficient Complete Solutions for Distributed Multi-Event Scheduling, in *AAAMAS04*, pp. 310–317 (2004)

[Mailler 04] Mailler, R. and Lesser, V.: Solving distributed constraint optimization problems using cooperative mediation, in *AAMAS04*, pp. 438–445 (2004)

[Miller 12] Miller, S., Ramchurn, S. D., and Rogers, A.: Optimal decentralised dispatch of embedded generation in the smart grid, in *AAMAS12*, Vol. 1, pp. 281–288 (2012)

[Modi 05] Modi, P. J., Shen, W., Tambe, M., and Yokoo, M.: Adopt: Asynchronous distributed constraint optimization with quality guarantees, *Artificial Intelligence*, Vol. 161, No. 1-2, pp. 149–180 (2005)

[Nguyen] Nguyen, D. T., Yeoh, W., and Lau, H. C.: Distributed Gibbs: A Memory-bounded Sampling-based DCOP Algorithm, in *AAMAS13*

[Okimoto 11] Okimoto, T., Joe, Y., Iwasaki, A., Yokoo, M., and Faltings, B.: Pseudo-Tree-Based Incomplete Algorithm for Distributed Constraint Optimization with Quality Bounds., in *CP11*, Vol. 6876 of *Lecture Notes in Computer Science*, pp. 660–674 (2011)

[Ottens 12] Ottens, B., Dimitrakakis, C., and Faltings, B.: DUCT: An Upper Confidence Bound Approach to Distributed Constraint Optimization Problems, in *AAAI12*, pp. 528–533 (2012)

[Petcu 05] Petcu, A. and Faltings, B.: A Scalable Method for Multiagent Constraint Optimization, in *IJCAI05*, pp. 266–271 (2005)

[Ramchurn 10] Ramchurn, S. D., Farinelli, A., Macarthur, K. S., and Jennings, N. R.: Decentralized Coordination in RoboCup Rescue, *Comput. J.*, Vol. 53, No. 9, pp. 1447–1461 (2010)

[Rogers 11] Rogers, A., Farinelli, A., Stranders, R., and Jennings, N. R.: Bounded Approximate Decentralised Coordination via the Max-Sum Algorithm, *Artificial Intelligence*, Vol. 175, No. 2, pp. 730–759 (2011)

[Vinyals 11] Vinyals, M., Shieh, E., Cerquides, J., Rodriguez-Aguilar, J. A., Yin, Z., Tambe, M., and Bowring, E.: Quality guarantees for region optimal DCOP algorithms, in *AAMAS11*, pp. 133–140 (2011)

[Zhang 05] Zhang, W., Wang, G., Xing, Z., and Wittenburg, L.: Distributed stochastic search and distributed breakout: properties, comparison and applications to constraint optimization problems in sensor networks, *Artificial Intelligence*, Vol. 161, No. 1-2, pp. 55–87 (2005)

[Zivan 08] Zivan, R.: Anytime Local Search for Distributed Constraint Optimization, in *AAAI08*, pp. 393–398 (2008)