Belief Propagationを取り入れた確率的関数進化アルゴリズム

東京大学大学院工学系研究科電気系工学専攻　伊庭研究室　佐藤浩之 伊庭斉志

GP-EDA (PMBGP)

- 木構造を対象としたEDA
- GPのオペレータの適用を確率分布の推定とサンプリングに置き換えたもの

1. プロトタイプ木　木構造を固定長1次元配列に変換しEDAを適用
2. 確率文脈自由文法　木構造を文脈自由文法で表現し適用確率を学習

Belief Propagation

- ノードがメッセージ(同時確率や周辺確率を計算するためのパーツ）を送り合ってグラフィカルモデル上での推論を効率的に行う
- グラフにループがないと必ず正しく有限のステップで停止
- ループがあると正しい結果に収束するとは限らない
- ループがある場合でも実用では良い成果(Loopy Belief Propagation)

1. max-sum (max-product)　最大同時確率と各ノードのインスタンスを推論
2. Sum-product　周辺確率と各ノードのインスタンスを推論

先行研究と研究の動機

１次元配列を対象としたEDAへのBelief Propagationへの導入は一定の成果を示している。そこで木構造を対象としたEDAにも有効に働くはずであると考え、今回の研究を行なった。

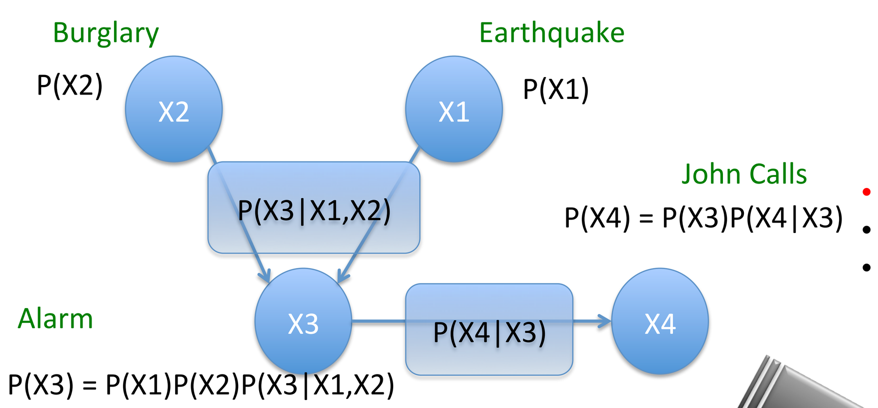
提案手法

- ➢ プロトタイプ木によるGP-EDA　POLE（Program Optimization with Linkage Estimation）
  - ➢ 確率モデル　ベイジアンネットワーク
  - ➢ 染色体　拡張構文木
- ➢ ファクターグラフ上でのloopy max-sum

ベイジアンネットワークを用いたGP-EDA	***提案手法***
① ベイジアンネットワークの構築 ② M個の個体をサンプリング	① ベイジアンネットワークの構築 ② M−1個の個体をサンプリング ③ ベイジアンネットワークをファクターグラフに変換 ④ ファクターグラフ上でloopy max-sumを走らせる ⑤ 最も同時確率の高い個体を生成

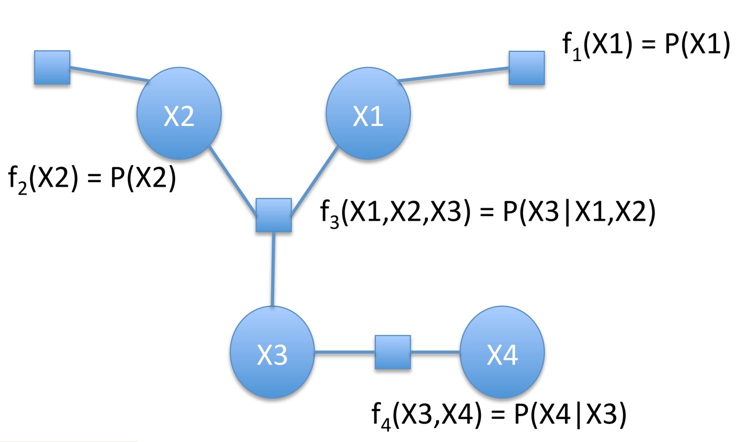
ベイジアンネットワークからファクターグラフへの変換

- ❖ どちらも図式的に表現される確率モデルであるグラフィカルモデルの1種
- ❖ ファクターグラフはベイジアンネットワークをより一般化したもの
- ❖ ベイジアンネットワークからファクターグラフの変換は常にできるが、逆はできるとは限らない



ベイジアンネットワーク

- 有向グラフ
- ノードは1種類
- エッジが条件付き確率を表す



ファクターグラフ

- 無向グラフ
- ノードは変数ノード、ファクターノードの2種類
- ファクターノードが確率と条件付き確率を表す

Loopy max-sum

1. 全てのメッセージを0で初期化
2. ノード同士で一定回数メッセージを送り合う

メッセージは同時確率の定義式を変形して導かれる。アンダーフローを防ぐために積を対数の和に置き換えている。

葉ノードから　$\mu_{x\to f}(x) = 0$　$\mu_{f\to x}(x) = \ln f(x)$

その他のノードから　$\mu_{f\to x}(x) = \max_{x,x_1,\ldots,x_M}\left[\ln f(x,x_1,\ldots x_M) + \sum_{m\in ne(f_s)\backslash x} \mu_{x_m\to f}(x_m)\right]$　$\mu_{x\to f}(x) = \alpha_{xf} + \sum_{l\in ne(x)\backslash f} \mu_{f_l\to x}(x)$

同時確率最大の個体を生成

同時確率を最大にする各ノードのインスタンスをとる

$$x^{\max} = \operatorname*{argmax}_x \left[\sum_{s\in ne(x)} \mu_{f_s\to x}(x)\right]$$

同時確率（対数スケール）

実験

DMAX問題

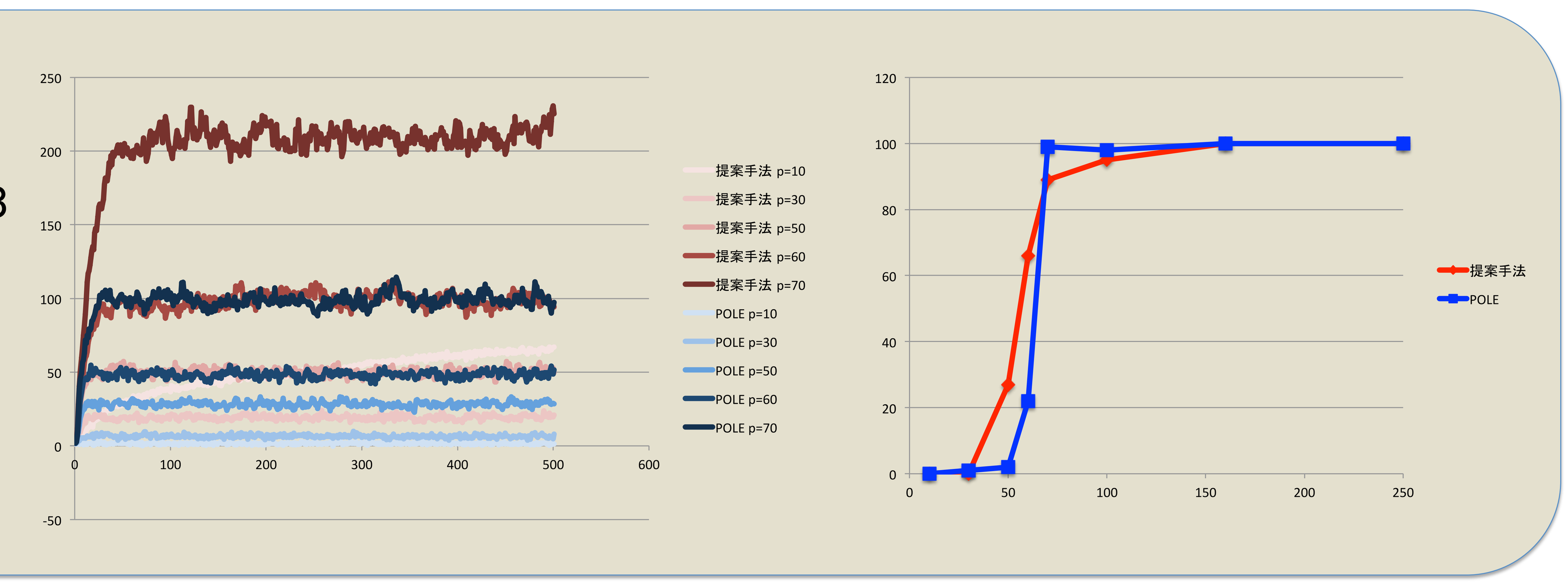
- ◆ MAX問題に騙し要素を加えたもの
- ◆ 目的は深さ制限の元で最大の実数成分を持つ関数を見つけること
- ◆ 終端ノードが複素数

$\mathfrak{F} = \{add_m, multiply_m\}$
$\mathfrak{T} = \{\lambda, 0.95\}$
$(\lambda)^r = 1, \lambda \in \mathbb{C}, r \in \mathbb{R}$

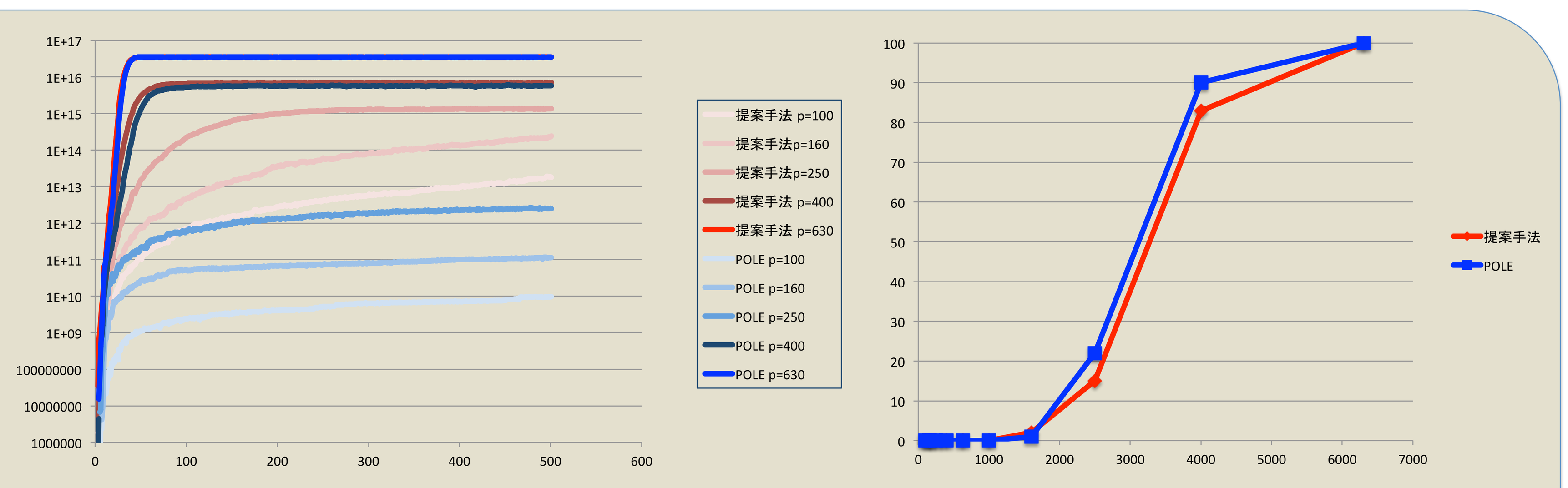
$add_m(a_0, a_1, \ldots, a_{m-1}) = \sum_{i=0}^{m-1} a_i$
$multiply_m(a_0, a_1, \ldots, a_{m-1}) = \prod_{i=0}^{m-1} a_i$

世代と平均適合度の推移　　100試行中何回最適解を獲得したか

深さ3



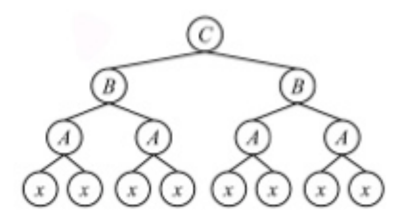
深さ4



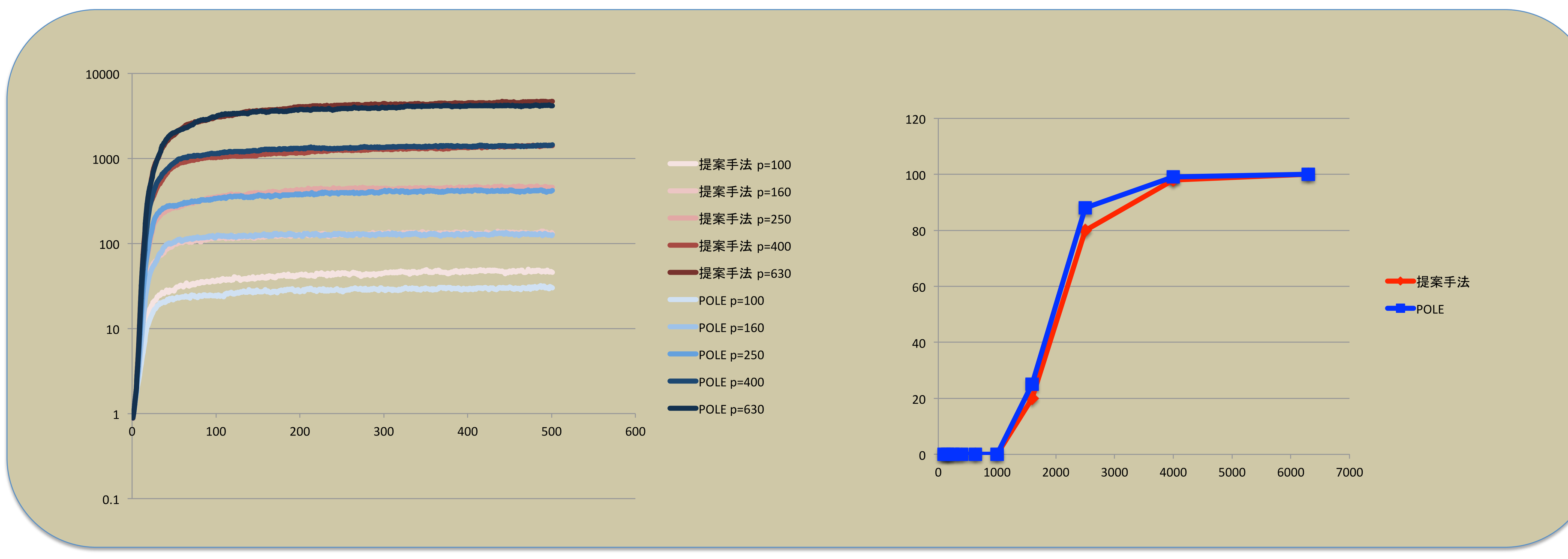
Royal Tree問題

- ◆ 関数ノードはA～Fのアルファベット、終端ノードはx(=1)とy(=0.95)
- ◆ アルファベットがひとつ前のアルファベットを子に持つ状態（Perfect Tree）が最適解
- ◆ 今回は深さ7で実験を行った

深さ3の最適解



世代と平均適合度の推移　　100試行中何回最適解を獲得したか



まとめと今後の課題

■ 個体数が少ない場合の平均適合度は提案手法のほうがよい。これはloopy max-sumが推定が十分でない確率モデルからも良い解を生成しやすいからであると考えられる。
■ 提案手法は従来の手法より悪くなることはないと想定していた。適合度上位の個体のみがベイジアンネットワークの構築に用いられるからである。しかし実際には最適解を獲得する割合は提案手法のほうが低かった。Loopy max-sumにより生成された適合度の高い個体が悪い方向に働くことがあると考えられる。
■　今回は騙し要素が強く解くことが難しい問題をベンチマークとして用いた。今後は簡単な問題やWall FollowingやRegressionといった性質の異なる問題に対しても性能評価を行いたい。
■　loopy max-sumは近似推論であり結果が正しいとは限らない。厳密推論であるJuction Treeの採用も考えている。

参考文献

1. Bishop, C. M.: Pattern Recognition and Machine Learning, Springer (2006).
2. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kauf- mann (1988).
3. Mendiburu, A., Santana, R., Lozano, J. A. and Bengoetxea", E.: A parallel framework for loopy belief propagation, in GECCO (Companion)'07, pp. 2843–2850 (2007)
4. 長谷川禎彦, 伊庭斉志:ベイジアンネットワーク推定による確率モデル遺伝的プログラミング, 人工知能学会論文誌, Vol. 22, No. 1, pp. 37–47 (2007).