2Q4-J-2-02

Augmented Naive Bayes Classifierの大規模構造学習

Learning huge Augmented Naive Bayes Classifier

菊谷成慎 Naruchika Kikuya　菅原聖太 Shouta Sugahara　名取和樹 Kazuki Natori　本田和雅 Kazunori Honda　植野真臣 Maomi Ueno

電気通信大学大学院情報理工学研究科
Graduate school of Informatics and Engineering, The University of Electro-Communications

For classification problems, Bayesian networks are often used to infer a class variable when given feature variables. Earlier reports have described that classification accuracies of exact learning augmented naive Bayes (ANB) achieved by maximizing the marginal likelihood (ML) were higher than the Bayesian network of the identification model However, the method cannot learn structures that have more than several dozen variables. To resolve this difficulty, this study proposed exact learning ANB using RAI algolithm. The experimental results show that the proposed method outperforms the other methods.

1. まえがき

ベイジアンネットワークは，離散確率変数をノードで表しノード間の依存関係を非循環有向グラフ（Directed Acyclic Graph: DAG）で表現する確率的グラフィカルモデルである．ベイジアンネットワークは，確率構造に DAG を仮定することにより，同時確率分布を条件付き確率の積に分解する．ベイジアンネットワークの構造は一般にデータから推定する必要があり，これをベイジアンネットワークの構造学習という．構造学習法として，漸近一致性を有する学習スコアを用いて，候補構造から最適なスコアを持つ構造を探索する厳密解探索アプローチが従来から用いられている．また，一般に学習スコアとして周辺尤度（Marginal Likelihood: ML）が用いられてきた．

ベイジアンネットワークにおける一つのノードを目的変数とし，その他のノードを説明変数としたベイジアンネットワーク分類器（Bayesain Network Classifier: BNC）は，離散変数を扱う分類器として知られている [Friedman 97]．BNC は説明変数を所与とした目的変数の条件付き確率をモデル化する識別モデルのほうが，通常のベイジアンネットワークを構造学習した生成モデルよりも分類精度が高いことが報告されている [Carvalho 11]．しかし，近年，Sugahara ら [Sugahara 18] は，データが十分大きいならば，生成モデルのほうが分類精度が高いことを示した．しかし，データが少ないときに目的変数の親が多い構造をとる場合，分類精度が著しく低下してしまうと指摘した．この問題を解決するために，彼らは目的変数が親を持たず，全ての説明変数を子に持つ構造を仮定した Augmented Naive Bayes（ANB）[Friedman 97] を厳密学習する手法を提案している．これにより，データが少ない場合も高い分類精度を得られることを経験的に示した．

しかし，厳密解探索アプローチによる厳密学習は NP 困難問題である．様々な学習アルゴリズムが提案されているが，未だに数十変数程度の構造学習が限界であり，Sugahara らの手法は多くの変数に対応できない．

一方，因果モデル分野では，漸近一致性は持たないが，より計算効率の高い制約ベースの構造学習法が提案されている．条件付き独立性検定（Conditional Independence test: CI テスト）と，オリエンテーションルールによるエッジの方向付けを行うことで DAG を効率的に学習する．従来の CI テストでは漸近一致性を持たないことが問題であったが，名取らは，最先端手法である RAI アルゴリズム [Yehezkel 09] の CI テストに Bayes factor を用いることで漸近一致性を有しつつ，1000 変数以上の大規模構造学習を実現している [名取 18]．さらに，本田らは推移性を組み込むことで，2000 変数程度の大規模構造学習を実現している [本田 19].

そこで本論文では，本田らの手法を ANB 学習に拡張させ，従来よりも大規模な BNC を学習することを目指す．さらに，評価実験により，提案手法が従来手法よりも大規模な説明変数を持つ分類問題で精度が高いことを示す．

連絡先: 電気通信大学大学院情報理工学研究科，調布市調布ヶ丘 1-5-1，042-484-8585，
{kikuya,sugahara,natori,hondak,ueno}@ai.lab.uec.ac.jp

2. ベイジアンネットワーク

ベイジアンネットワークは，離散確率変数をノードとし，ノード間の依存関係を非循環有効グラフ（Directed Acyclic Graph: DAG）と各ノードの条件付き確率で表現する確率的グラフィカルモデルである．今，$\mathbf{X} = \{X_0, X_1, \ldots, X_n\}$ を離散確率変数集合とし，各変数 X_i は r_i 個の状態集合 $\{1, \ldots, r_i\}$ から一つの値 k を取る（$X_i = k$ と書く）とする．このとき，構造 G において，各ノード X_i の親ノード集合を Π_i としたときの同時確率分布 $P(X_0, X_1, \ldots, X_n \mid G, \Theta)$ は以下のように表現できる．

$$P(X_0, X_1, \ldots, X_n \mid G, \Theta) = \prod_{i=0}^{n} P(X_i \mid \Pi_i, G, \Theta). \tag{1}$$

ここで，θ_{ijk} を Π_i が j 番目のパターンを取るとき ($\Pi_i = j$ と書く) に $X_i = k$ となる条件付き確率 $P(X_i = k \mid \Pi_i = j, G)$ を示すパラメータとし，パラメータ集合 $\Theta = \{\theta_{ijk}\}, (i = 0, \ldots, n; j = 1, \ldots, q_i; k = 1, \ldots, r_i)$ とする．

ベイジアンネットワークでは，パラメータ推定値として，期待事後確率推定値（Expected a Posteriori: EAP）が最もよく用いられる．変数集合 $\mathbf{X}$ に対する N 個のデータを $\mathbf{D} = \{D_1, \ldots, D_N\}$ とすると，EAP は事前分布にディリクレ分布を仮定することで次式で求まる．

$$\hat{\theta}_{ijk} = \frac{\alpha_{ijk} + N_{ijk}}{\alpha_{ij} + N_{ij}}. \tag{2}$$

ここで，N_{ijk} は，X_i の親変数集合 Π_i が j 番目のパターンを取った時の $X_i = k$ となるデータの頻度を表し，$n_{ij} = \sum_{k=1}^{r_i} n_{ijk}$ を表す．また，$\boldsymbol{\alpha} = \{\alpha_{ijk}\}, (i = 0, \ldots, n; j = 1, \ldots, q_i; k = 1, \ldots, r_i)$ はディリクレ事前分布のハイパーパラメータであり，$\alpha_{ij} = \sum_{k=1}^{r_i} \alpha_{ijk}$ である．

ベイジアンネットワークは最適な構造もデータから推定する必要がある．これをベイジアンネットワークの構造学習という．構造学習法として，候補構造から最適なスコアを持つ構造を探索する厳密解探索アプローチが従来から用いられている．一般に学習スコアとして周辺尤度 $P(\mathbf{D} \mid G)$（Marginal Likelihood: ML）が用いられる．ML はパラメータの事前分布にディリクレ分布を仮定することで次のように閉形式で表せる．

$$P(\mathbf{D} \mid G) = \prod_{i=0}^{n} \prod_{j=1}^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_{ij} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk} + N_{ijk})}{\Gamma(\alpha_{ijk})}. \quad (3)$$

近年では，$\alpha_{ijk} = \alpha/(r_i q_i)$ とした Bayesian Dirichlet equivalent uniform (BDeu) が最も用いられる．ここで，α は Equivalent Sample Size (ESS) と呼ばれる事前知識の重みを示す疑似サンプルである．BDeu を用いた厳密学習は漸近一致性を持つことから，学習データが十分大きいならば真の構造を推定することが知られている．

しかし，厳密解探索アプローチは NP 困難であり，変数数に対して計算時間が爆発的に増加してしまう．厳密解を効率的に探索するために，動的計画法，A*探索，整数計画法などが提案されているが，数十変数の構造学習が限界であり，大規模構造を学習することができない．

3. ベイジアンネットワーク分類器

ベイジアンネットワーク分類器（BNC）は，ベイジアンネットワークにおける一つのノードを目的変数，それ以外のノードを説明変数とした分類器であり，高い分類精度を持つ [Friedman 97]．今，X_0 を目的変数とし，$X_1, \ldots, X_n$ を説明変数とした BNC を考える．説明変数のデータ $\mathbf{e} = \langle x_1, \ldots, x_n \rangle$ が与えられた時，目的変数の推定値 $\hat{c}$ は以下のように得られる．

$$\begin{aligned}
\hat{c} &= \underset{c \in \{1,\ldots,r_0\}}{\arg\max}\ P(c \mid x_1, \ldots, x_n, G) \quad (4)\\
&= \underset{c \in \{1,\ldots,r_0\}}{\arg\max} \prod_{i=0}^{n} \prod_{j=1}^{q_i} \prod_{k=1}^{r_i} (\theta_{ijk})^{1_{ijk}}\\
&= \underset{c \in \{1,\ldots,r_0\}}{\arg\max} \left[\prod_{j=1}^{q_0} \prod_{k=1}^{r_0} (\theta_{0jk})^{1_{0jk}} \right.\\
&\qquad \left. \times \prod_{i: X_i \in \mathbf{C}} \prod_{j=1}^{q_i} \prod_{k=1}^{r_i} (\theta_{ijk})^{1_{ijk}} \right]
\end{aligned}$$

ここで，1_{ijk} はデータ $\mathbf{e}$ が与えられた時，$X_i = k$ かつ $\Pi_i = j$ の時 1 をとり，それ以外の場合は 0 をとる変数である．さらに，$\mathbf{C}$ は目的変数の子集合である．式 (4) より，目的変数 X_0 の親集合，子集合，X_0 と子を共有している変数集合の和集合のみで推定値 $\hat{c}$ を求めることができる．また，この集合を X_0 のマルコフブランケット（Markov Blanket: MB）と呼ぶ．

BNC の構造学習は通常，ベイジアンネットワークと同様にデータから学習する．構造に制約を持たない BNC を General Bayesian Network（GBN）と呼ぶ．しかし，BDeu などで学習された GBN は目的変数が子を持たず，親を多く持つ構造をとることがある．このような構造では条件付きパラメータ数が多くなり，パラメータの推定に用いられるデータがスパースになる．その結果，パラメータの推定精度が悪くなり，分類精度が著しく低くなる傾向がある．この問題を回避できる BNC として，目的変数が全ての説明変数を子に持ち，説明変数間が独立であると仮定する Naive Bayes，目的変数が全ての説明変数を子に持ち，説明変数間で木構造をとると仮定した Tree-Augmented Naive Bayes（TAN）[Friedman 97] などが提案されている．また，これらを一般化したモデルとして，目的変数が全ての説明変数を子に持つことのみ仮定する Augmented Naive Bayes（ANB）[Friedman 97] が知られている．

ML を最大にする BNC は同時確率をモデル化した生成モデルであるが，説明変数を所与とした目的変数の条件付き尤度をモデル化する識別モデルのほうが，漸近的な分類精度が高いことが報告されていた [Carvalho 13]．しかし，識別モデルとしての BNC が生成モデルとしての BNC よりも高い分類精度を得られる根拠が理論的に示されていない．そこで，Sugahara ら [Sugahara 18] は ML で厳密学習した BNC は，識別モデルの BNC よりも生成モデルの分類精度が低いとは限らないことを示した．しかし，先述したように生成モデルの BNC は目的変数の親が多い構造を学習した場合，分類精度が著しく低下してしまう．この問題を解決するため，彼らは ANB を生成モデルとして厳密学習する手法を提案した．結果は，提案手法がデータが少ない場合でも安定した分類精度を得ることができ，識別モデルの BNC より有意に分類精度が高いことを示した．しかし，厳密学習は数十変数の学習が限界であり，変数数が多いデータに対して用いることができない．そこで本論文では，BNC の大規模構造学習法を提案する．

4. 制約ベース厳密構造学習

4.1 Bayes factor を用いた制約ベース厳密学習

制約ベースアプローチは計算量を大幅に削減できる構造学習法である．このアプローチの基本的なアルゴリズムは次のとおりである．

(1) 完全無向グラフを生成する．

(2) (1) で生成された完全無向グラフに対し条件付き独立検定（Conditional Independence test: CI テスト）によりエッジを削除する．

(3) (2) で得られた無向グラフに対してオリエンテーションルールを用いて方向付けを行う．

一般に，制約ベースアプローチの学習精度は CI テストの精度に依存し，学習速度は学習に要する CI テストの回数に依存する．制約ベースアプローチとして，PC アルゴリズム，MMHC アルゴリズム，RAI アルゴリズム [Yehezkel 09] が提案されてきた．しかし，これらのアルゴリズムでは χ^2 検定，G^2 検定，条件付き相互情報量などを CI テストに用いるため，漸近一致性を持たない．

一方で，二変数間が従属・独立モデルの周辺尤度の比による Bayes factor を用いた漸近一致性を有する CI テストが提案されている．ベイジアンネットワークのノード X,Y において X と Y 間について各ノードの共通親ノード集合 $\mathbf{Z}$ をとしたときの従属なモデルを G_1，独立モデルを G_2 としたときの Bayes factor を $\mathrm{BF}(X, Y \mid \mathbf{Z})$ とすると，対数 Bayes factor は，

$$\log \mathrm{BF}(X, Y \mid \mathbf{Z}) = \log \frac{P(\mathbf{D} \mid G_1, \boldsymbol{\alpha})}{P(\mathbf{D} \mid G_2, \boldsymbol{\alpha})} \quad (5)$$

と表される．ここで，$P(\mathbf{D} \mid G_1, \boldsymbol{\alpha})$，$P(\mathbf{D} \mid G_2, \boldsymbol{\alpha})$ は式（3）の BDeu を用いる．Bayes factor を用いた CI テストでは，対数 Bayes factor が 0 以上か否かで従属，独立を判断する．名取ら [名取 18] は Bayes factor を用いた CI テストが漸近的に真の条件付き独立性を判定できることを示した．さらに，Bayes factor を用いた CI テストを RAI アルゴリズムに組み込んだ手法を提案し，漸近一致性を有しつつ 1000 変数を超える大規模構造学習を実現した．

4.2 推移性を用いた厳密アルゴリズム

前述のように、制約ベースの学習速度はＣＩテストの回数に依存する．一般にＣＩテストの回数は学習のできる限り早期にエッジを削除するほど削減できる．そのために以下の推移性を用いる手法が提案されている．

定理 4..1 $G = (\boldsymbol{V}, \boldsymbol{E})$ を DAG とし，$X, Y \in \boldsymbol{V}$ で，Y は X の非子孫とする．このとき，$A \in \boldsymbol{V} \backslash (\{X, Y\} \cup \boldsymbol{Pa}(X, G) \cup \boldsymbol{W})$ とすると，以下が成り立つ．

$$X \perp Y \mid \mathbf{Pa}(X, G) \rightarrow X \perp A \mid \mathbf{Pa}(X, G) \; or \; A \perp Y \mid \mathbf{Pa}(X, G) \tag{6}$$

ここで，$\mathbf{W}$ は X の子孫であり，X と Y が合流結合するノードとその子孫からなるノード集合を表し，$\mathbf{Pa}(X, G)$ は X の G における親ノード集合を表す．証明は [本田 19] を参照してほしい．定理 4..1 より，ある二変数の条件付き独立性からその各変数と他変数との条件付き独立性の少なくとも一つを保証できる．これを利用することで，二変数の条件付き独立性から少なくとも一つの条件付き独立が保証される二組のエッジの CI テストを列挙できる．列挙されたＣＩテストを優先して行うことでより学習の早期にエッジを削除できる．本田らは推移性によるエッジの削除法を Bayes factor を用いた RAI アルゴリズムに組み込むことで CI テストの回数を削減し，2000 変数程度の大規模構造を実現した．

本論文では，本田らの手法 [本田 19] を用いることで，従来よりも大規模な BNC の構造学習の実現を目指す．

5. 提案手法

本章では，本田らの手法を ANB 学習に拡張させる方法を示す．まず，RAI アルゴリズムの基本的な動作を紹介する．今，グラフを $G = (\mathbf{V}, \mathbf{E})$ と表し，$\mathbf{V}, \mathbf{E}$ はそれぞれ G に含まれるノード集合，エッジ集合を表す．ここで，G は有向エッジと無向エッジを併せ持つとする．また，$G_{ex} = (\mathbf{V}_{ex}, \mathbf{E}_{ex})$ を RAI アルゴリズムによって分割された部分グラフとする．

(1) 完全無向グラフ G_{uc} とデータ $\mathbf{D}$ を入力する．

(2) 各次数の CI テストにおいて $\log \mathrm{BF}(X, Y \mid \mathbf{Z}) < 0$ となり X，Y が条件付き独立と判定されるとき，XY 間のエッジを削除する．

(3) 1 次以降の CI テストによるエッジの削除の直後，推移性に基づくエッジの削除を行う．ここでは，(2) の CI テストで推定した条件付き独立性から推移性により条件付き独立が成り立つ候補ノード集合 $\mathbf{A}$ を検出する．その後，X と A，A と Y（$A \in \mathbf{A}$）に対してそれぞれ CI テストを行い，条件付き独立と判定されるときエッジを削除する．

(4) 得られたグラフに対してオリエンテーションルールを適用して方向付けを行う．

(5) 方向付けの結果から部分グラフ G_{ex} に分割する．

(6) 各部分グラフで再帰的に RAI を呼び出す．

アルゴリズムの詳細は，[名取 18] や [本田 19] を参照してほしい．

ANB は目的変数が親を持たず，全ての説明変数を子に持つ構造である．上記のアルゴリズムで推定される構造に ANB を仮定するために，手順 (1) の初期グラフと，手順 (2)，手順 (3) の CI テストの実施範囲に制約を加えることで，RAI アルゴリズムを用いた ABN 厳密学習を実現する．ここで，目的変数のノードを X_0 それ以外のノードを目的変数のノードとする．

まず手順 (1) において，初期グラフを完全無向グラフ G_{uc} ではなく，G_{uc} に対して，目的変数 X_0 から全ての説明変数に向けてエッジの方向付けを行う．つまり，目的変数から説明変数へ向かう有向エッジと，説明変数間の無向エッジを併せ持つ完全グラフを初期グラフとする．

次に手順 (2) において，目的変数と説明変数を接続するエッジが CI テストで削除されないように，CI テストの実施範囲に制約を加える．つまり，目的変数 X_0 を除く二変数 $X', Y' \in \mathbf{V} \setminus \{X_0\}$ に対して CI テストを実施する．

手順 (3) においても同様に CI テストの実施範囲に制約を加える．ここでは，推移性により検出された候補ノード集合 $\mathbf{A}$ から目的変数 X_0 を除いた $A' \in \mathbf{A} \setminus \{X_0\}$ に対して CI テストを実施する．

これらの変更を加えることで RAI アルゴリズムは出力グラフに ANB を返す．

6. 評価実験

6.1 実験手順

UCI リポジトリデータベースから，変数数が大きな 13 個のデータセットを用いて従来手法で学習した BNC と提案手法で学習した BNC の分類精度を比較した．各データセットは欠損値を含まず，データに含まれる連続量はいずれも中央値を境に 2 値に離散化した．比較手法は，Naive Bayes，対数尤度を最適化した TAN（TAN），RAI アルゴリズムを用いて学習した GBN（RAI-GBN），RAI アルゴリズムを用いて学習した ANB（RAI-ANB）である．いずれの RAI アルゴリズムも CI テストに Bayes factor（ESS=1.0）を適用している．2 章で述べたように厳密解探索アプローチは数十ノードの学習が限界であるため比較対象から除外した．本論文では，RAI-ANB を提案手法とする．

また，解析の妥当性を示すため，各 BNC の構造学習と分類に対して 10 分割交差検証を行った．表 1 に各データセットに対する各手法の分類精度を載せた．提案手法の有意性を示すため，各手法の分類精度に対し Hommel 法を用いて多重検定を行い，p 値を表 1 の最下部に載せた．

6.2 結果と考察

表 1 より，提案手法は他手法よりも有意水準 0.05 のもとで有意に分類精度が高かった．また，RAI-GBN は分類精度の平均が最も低かった．表 2 は各データセットにおける RAI-GBN で学習した構造のエッジ数，目的変数の親と子の数，マルコフブランケット（MB）の平均を表している．表 2 より，RAI-GBN は，ほとんどのデータセットに対して親が多く，子が少ない構造を学習している．また，MB の数が全説明変数に対して小さ

表 1: 提案手法と従来手法の分類精度（太字は最大の分類精度）

	データセット	説明変数	目的変数の状態数	データ数	Naive Bayes	TAN	RAI-GBN	RAI-ANB
1	Connect-4	42	3	67557	0.7213	0.7643	0.7337	**0.7928**
2	dota2	116	2	102944	**0.5981**	0.5810	0.5442	0.5957
3	Epileptic Seizure	178	5	11500	0.2344	0.3650	0.1887	**0.4044**
4	Flowmeters D	43	4	180	**0.8389**	0.8389	0.6778	0.8278
5	kr-vs-kp	36	2	3196	0.8774	0.9240	0.9406	**0.9468**
6	madelon	500	2	2000	0.5905	0.5270	**0.6215**	0.5820
7	mfeat-fac	218	10	2000	0.3520	0.4590	0.2630	**0.4610**
8	MicroMass	1300	10	360	0.9472	0.9472	0.7361	**0.9500**
9	movement libras	90	15	360	0.5028	0.5389	0.2278	**0.5583**
10	Musk1	166	2	478	0.6517	0.7566	0.6744	**0.7965**
11	Musk2	166	2	6598	0.7445	0.8406	0.8821	**0.9627**
12	Parkinson's Disease	754	2	756	0.7182	0.7898	0.7672	**0.8108**
13	semeion	256	10	1600	0.8550	0.8719	0.4557	**0.8745**
	平均				0.6640	0.7080	0.6064	**0.7356**
	p 値				0.0026	0.0012	0.0009	-

表 2: RAI-GBN の統計概要

	データセット	エッジ数	子変数	親変数	MB
1	Connect-4	92.6	8.4	1.7	17.7
2	dota2	151.6	0	3.0	3.0
3	Epileptic Seizure	325.1	0	0	0.0
4	Flowmeters D	63.6	0	3.7	3.7
5	kr-vs-kp	249.4	0	5.1	5.1
6	madelon	249.4	0.1	2.6	3.0
7	mfeat-fac	630.7	0	3.5	3.5
8	MicroMass	2892.4	0	7.0	7.0
9	movement libras	109.7	0	2.4	2.4
10	Musk1	413.2	0	2.0	2.0
11	Musk2	946.9	0	6.1	6.1
12	Parkinson's Disease	1475.6	0	2.4	2.4
13	semeion	820.1	0	4.0	4.0

いことから，ほとんどの説明変数が目的変数の推定に関与していないことがわかる．また，目的変数の状態が多いデータを用いた場合，他手法と比べて著しく分類精度が低いことがわかる．このことから，データ数が不十分である場合，厳密解探索アプローチによる厳密学習の場合と同様に分類精度が著しく低くなることがわかる．それに対し提案手法は，目的変数が全ての説明変数を子に持つため，全ての説明変数が目的変数の分類に寄与している．また，目的変数が親を持たないため，パラメータ数が抑えられ，データ不足が緩和される．これらの理由から，提案手法を用いることで分類精度が高い構造が学習できたと考えられる．

7. むすび

本論文では，ベイズファクターを用いた RAI アルゴリズムを ANB 学習に拡張させることで，従来より大規模で高精度な BNC を学習する手法を提案した．提案手法は変数数が 1000 以上の大規模な ABN を学習できることを示した．また，評価実験により，提案手法は同規模の構造学習が可能な従来の BNC よりも分類精度が有意に高いことを示した．

参考文献

[Carvalho 11] Carvalho, A. M., Roos, T., Oliveira, A. L., and Myllymäki, P.: Discriminative Learning of Bayesian Networks via Factorized Conditional Log-Likelihood, *Journal of Machine Learning Research*, Vol. 12, pp. 2181–2210 (2011)

[Carvalho 13] Carvalho, A. M., Adão, P., and Mateus, P.: Efficient Approximation of the Conditional Relative Entropy with Applications to Discriminative Learning of Bayesian Network Classifiers, *Entropy*, Vol. 15, No. 7, pp. 2716–2735 (2013)

[Friedman 97] Friedman, N., Geiger, D., and Goldszmidt, M.: Bayesian Network Classifiers, *Machine Learning*, Vol. 29, No. 2, pp. 131–163 (1997)

[Sugahara 18] Sugahara, S., Uto, M., and Ueno, M.: Exact learning augmented naive Bayes classifier, in *International Conference on Probabilistic Graphical Models*, Vol. 72, pp. 439–450 (2018)

[Yehezkel 09] Yehezkel, R. and Lerner, B.: Bayesian network structure learning by recursive autonomy identification, *Journal of Machine Learning Research*, Vol. 10, pp. 1527–1570 (2009)

[本田 19] 本田和雅, 名取和樹, 菅原聖太, 磯崎隆司, 植野真臣：推移性を利用した大規模ベイジアンネットワーク構造学習, Master's thesis, 電気通信大学 (2019)

[名取 18] 名取和樹, 宇都雅輝, 植野真臣：Bayes factor を用いた RAI アルゴリズムによる大規模ベイジアンネットワーク学習, 電子情報通信学会論文誌 D, Vol. 101, (2018)