

ベイズ的人工知能特論 ガイダンス

植野真臣
電気通信大学
大学院情報システム学研究科

本日の目標

- 自己紹介
- 分野・コミュニティの紹介
- ベイジアンネットワークのものすごく大ざっぱな紹介
- 授業内容・方法・評価方法の説明

ビック・データ時代

- 90年代—00年代 インフラストラクチャー時代
いかにデータを蓄えるか、データを蓄えさせる時代
- 有り余るデータとその有効活用が課題:大量のデータから高精度に高度な処理ができる手法→次世代の企業コンピューテンシー
- 簡単にまねできない高精度な処理技法が注目される時代に突入
- 総合格闘技→数理統計、アルゴリズム、データベースの統合的技術

人工知能の簡単な歴史

- 古典的AIの破たん
論理推論、IF then rule
- 人は死ぬ、ソクラテスは人である、ソクラテスは死ぬ
.
- 古典的AIの問題
当たり前のことしか推論できない

最先端人工知能

- 確率推論
- 例 鳥は飛ぶ、ペンギンは鳥、ペンギンは飛ばない
- ロジック推論では例外処理が必要
- $P(\text{鳥は飛ぶ})=0.9$
- $P(\text{それは鳥} \mid \text{それは飛ばない, 他の証拠}) > 0$
- メリット
- 論理推論も確率1-0として表現できる！！

同時確率分布が世界を表現！！

- 性別、髪の長さ、背の高さ
 $P(\text{男、髪短い、背高い})=0.2$
 $P(\text{男、髪長い、背高い})=0.125$
 $P(\text{男、髪長い、背低い})=0.05$
 $P(\text{男、髪短い、背低い})=0.125$
 $P(\text{女、髪短い、背高い})=0.05$
 $P(\text{女、髪長い、背高い})=0.125$
 $P(\text{女、髪長い、背低い})=0.2$
 $P(\text{女、髪短い、背低い})=0.125$
- $P(\text{男})=0.5, P(\text{女})=0.5$

同時確率分布を知っていると

- 「その人は髪が短い」がわかった

$P(\text{男、髪短い、背高い})=0.2$

$P(\text{男、髪短い、背低い})=0.125$

$P(\text{女、髪短い、背高い})=0.05$

$P(\text{女、髪短い、背低い})=0.125$

男の確率= $0.325/0.5=0.65$

世界中のすべての変数の同時確率分布を知ればなんでも推論できる！！

問題

- 変数が増えると状態数が指数的に増える
- すべて二値しかとらなくても 10変数で1024の同時確率を推定しないとイケない。

二つの問題

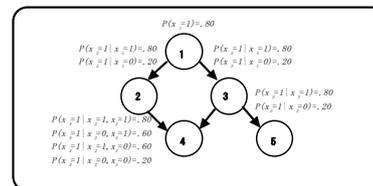
- 計算量が指数的に爆発する。
- データからの推定値がスパースになる。
(大量のデータがあっても空データだらけになる)

ベイジアンネットワークとは

- 確率構造が非循環有向グラフであれば、同時確率分布が条件付確率の積に因数分解できる数学的定理により、同時確率分布を表現する厳密解に極めて近い近似。
- 確率有向グラフが確率因果構造が対応し、ものごとの因果もわかる！！

現在考えられる最もよい同時確率分布の近似
→ 推論の予測精度が最高のはず！！

ベイジアン・ネットワークモデル

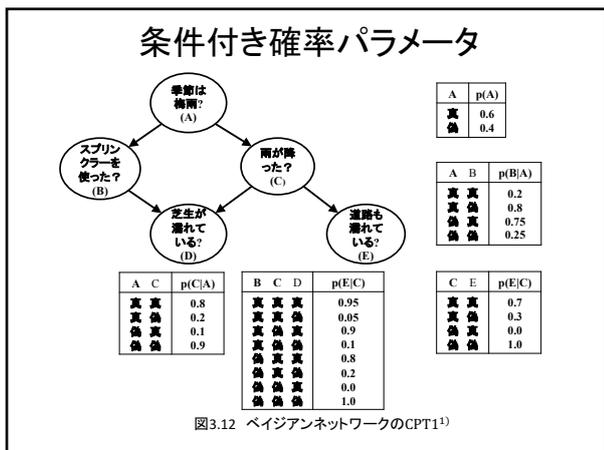


同時確率パラメータは32個

条件付き確率パラメータは11個

$$P(x_1, x_2, \dots, x_N | G) = \prod_{i=1}^N p(x_i | \Pi_i, G)$$

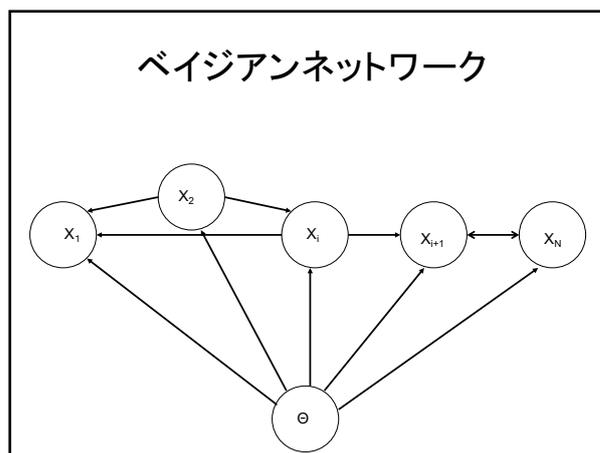
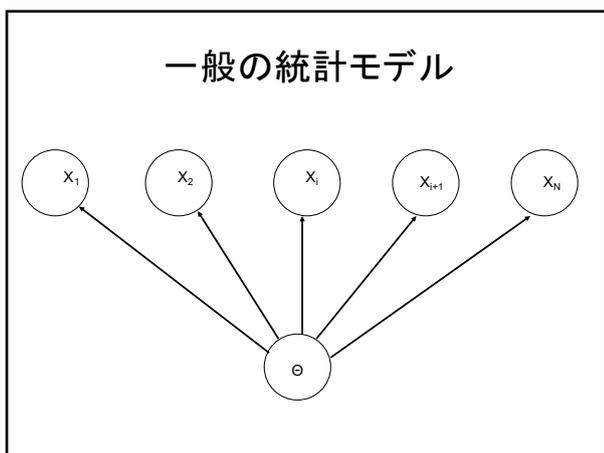
$\Pi_i \subseteq \{x_1, x_2, \dots, x_{q_i}\}$ は変数iの親ノード集合



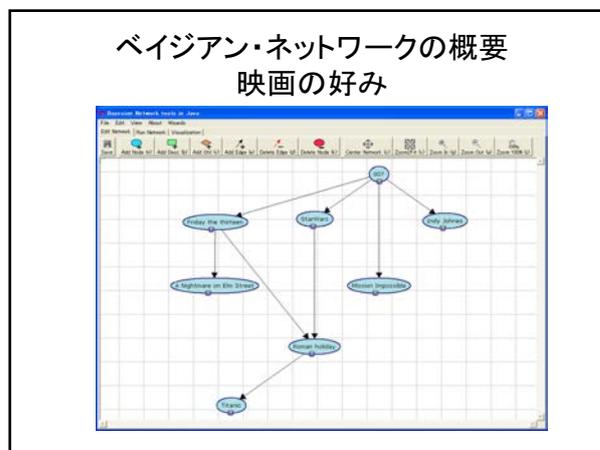
同時確率分布表: joint probability distribution table, JPDT

表3.1 図3.2の同時確率分布表:JPDT

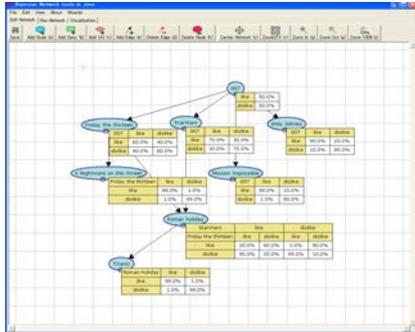
| | | | | | | | | | | | |
|---|---|---|---|---|---------|---|---|---|---|---|---------|
| 1 | 1 | 1 | 1 | 1 | 0.06384 | 0 | 1 | 1 | 1 | 1 | 0.01995 |
| 1 | 1 | 1 | 1 | 0 | 0.02736 | 0 | 1 | 1 | 1 | 0 | 0.00855 |
| 1 | 1 | 1 | 0 | 1 | 0.00336 | 0 | 1 | 1 | 0 | 1 | 0.00105 |
| 1 | 1 | 1 | 0 | 0 | 0.00144 | 0 | 1 | 1 | 0 | 0 | 0.00045 |
| 1 | 1 | 0 | 1 | 1 | 0.0 | 0 | 1 | 0 | 1 | 1 | 0.0 |
| 1 | 1 | 0 | 1 | 0 | 0.02160 | 0 | 1 | 0 | 1 | 0 | 0.24300 |
| 1 | 1 | 0 | 0 | 1 | 0.0 | 0 | 1 | 0 | 0 | 1 | 0.0 |
| 1 | 1 | 0 | 0 | 0 | 0.00240 | 0 | 1 | 0 | 0 | 0 | 0.02700 |
| 1 | 0 | 1 | 1 | 1 | 0.21504 | 0 | 0 | 1 | 1 | 1 | 0.00560 |
| 1 | 0 | 1 | 1 | 0 | 0.09216 | 0 | 0 | 1 | 1 | 0 | 0.00240 |
| 1 | 0 | 1 | 0 | 1 | 0.05376 | 0 | 0 | 1 | 0 | 1 | 0.00140 |
| 1 | 0 | 1 | 0 | 0 | 0.02304 | 0 | 0 | 1 | 0 | 0 | 0.00060 |
| 1 | 0 | 0 | 1 | 1 | 0.0 | 0 | 0 | 0 | 1 | 1 | 0.0 |
| 1 | 0 | 0 | 1 | 0 | 0.0 | 0 | 0 | 0 | 1 | 0 | 0.0 |
| 1 | 0 | 0 | 0 | 1 | 0.0 | 0 | 0 | 0 | 0 | 1 | 0.0 |
| 1 | 0 | 0 | 0 | 0 | 0.09600 | 0 | 0 | 0 | 0 | 0 | 0.09000 |



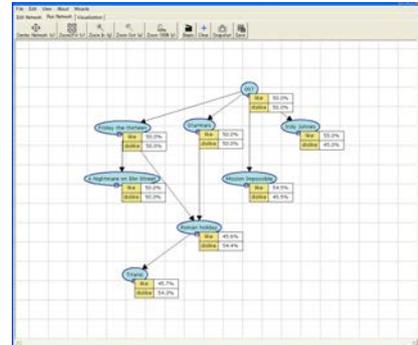
- ### ベイジアンネットワークの利点
- グラフと確率因果構造が対応し、因果関係が解釈しやすい
 - 仮定の少ないモデルによる同時確率分布の近似
 - ->
 - 多くの下位モデルを含んでいる。
 - 今考えられるもっとも一般化された同時確率分布の近似であり、数学的な意味がある。
 - 厳密な学習法と予測手法の存在->最高性能の予測率が期待できる。(最悪、推定された因果モデルが間違っていたとしても同時確率分布の近似はよく、予測についても精度が高い。)
 - 真のモデルがBNのクラスであれば、正しい因果モデルが得られる



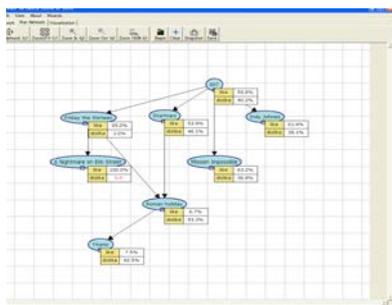
各アークに条件付確率を作成



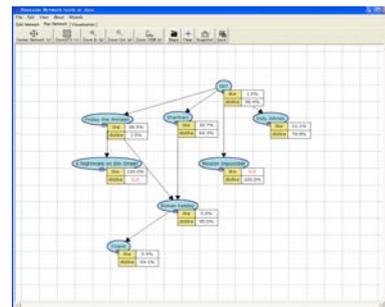
事前の確率



証拠「「エルム街の悪夢」が好き」



証拠「「ミッション・インポッシブル」は嫌い」



ベイジアンネットワークの主な研究課題

- 学習
- 推論

じゃ！！なぜみんなベイジアンネットワークを研究しないのか？

じゃ！！なぜみなベイジアンネットワークを研究しないの？

- 答え
- 難しいからです。

- だからベイジアンネットワークを学ぶことは競争力があるのです！！

ベイジアンネットワークソフト

- 商用ソフト
- Hugin
- BAYO NET
- フリーソフト
- Bayes Net Toolbox (BNT, <http://bnt.googlecode.com>)
- the Probabilistic Networks Library (PNL, see <http://sourceforge.net/projects/openpnl>)
- FastInf (Jaimovich et al., 2010), GRMM (<http://mallet.cs.umass.edu/grmm>)
- Factorie (<http://code.google.com/p/factorie>)
- JProGraM (<http://jprogram.sourceforge.net>)

しかし

問題

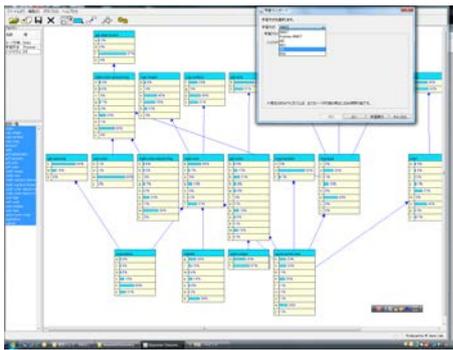
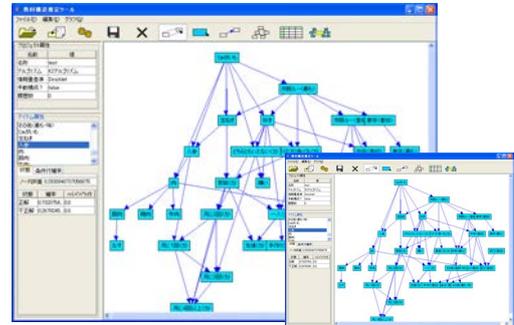
- ほとんどのソフトウェアは まだ「なんちゃって」ベイジアンネットワークソフトです！！
- 例
Hugin 推論は厳密解であるが、学習機能はいいかげん
BayoNet : 学習も推論も厳密ではない！！

これには大きな理由がある！！
ベイジアンネットワークの問題

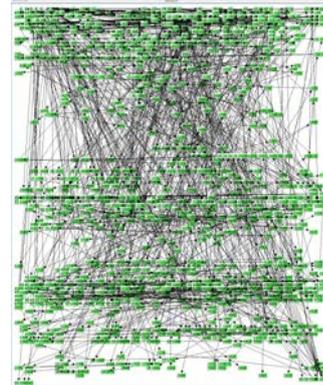
- 欠点として計算量の多さ(現在 厳密学習では、30-50ノードのネットワーク、厳密推論では80-100ノードのネットワーク)がある
- 将来的にはこれを克服すれば最強ツールになる！！

- Ueno UAI 2010, 2011
- 最もよいベイジアンネットワーク学習スコアの導出
- Natori 他 AMBN2015, PGM2016 学習できる変数数1000変数以上へ拡張
- Li and Ueno 2015 Approximate Reasoning 厳密推論 200以上の変数の推論へ拡張

植野研ベイジアン・ネットワーク



大規模ベイジアンネットワーク・エンジン

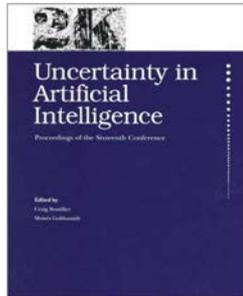


ベイジアンネットワークコミュニティ

まずは自己紹介！！

- 植野真臣 (うへのまおみ)
- 専門 人工知能、機械学習、確率推論、ベイズ統計
- 学会活動
 - AAAI Program Committee
 - IJCAI Program Committee
 - PGM Program Committee
 - AMBN General Chair
 - AIED Program Committee
 - IEEE ICTAI Best Paper Winner
 - など

UAI: ベイジアンネットワークの最高峰



Prof. Judea Pearl 2012 Turing Award Winner



David Heckerman : UAI President



Former chairman of the computer science department at **UCLA**

Abnan Darwiche is also a former Editor-in-Chief for the Journal of Artificial Intelligence Research (**JAIR**) and a **AAAI Fellow**



Prof. Millymaki: Helsinki大: コンピュータ科学 学科長



アルゴリズムの神: Knuth



ZDDによるBN推論

2004年 クヌースのZDDによるBN推論

数え上げ!!

ベイジアンネットワークに必要な技術

- 離散数学、数え上げ理論
- アルゴリズム、グラフ理論とデータベース
- 確率論、統計的学習理論、機械学習、統計物理
- ハイパフォーマンスコンピューティング

AMBN Since 2010



Alessandro Antonucci (IDSIA, Switzerland)
 Cassio P. de Campos (Queens University, UK)
 Hei Chan (SM)
 Arthur Choi (UCLA, UCLA)
 Adnan Darwiche (UCLA, USA)*
 Robin Evans (University of Oxford, UK)
 Luca Faes (University of Trento, Italy)
 David Heckerman (Microsoft, USA)*
 Antti Hyttinen (University of Helsinki, Finland)
 Aapo Hyvarinen (University of Helsinki, Finland)*
 Selya Imoto (University of Tokyo, Japan)
 Yoshinobu Kawahara (Osaka university)
 Manabu Kuroki (SM)
 Jose A. Lozano (University of the Basque Country UPV/EHU, Spain)
 Peter Lucas (Institute for Computing and Information Sciences, The Netherlands)
 Brandon Malone (Max Planck Institute, Germany)
 Alessio Moneta (Scuola Superiore Sant'Anna, Italy)
 Yoichi Motomura (National Institute of Advanced Industrial Science and Technology, Japan)
 Petri Myllymäki (University of Helsinki, Finland)*
 Judea Pearl (UCLA, USA)*
 Jose M. Pena (Linköping University)
 Hiroshi Sakamoto (Kyushu Institute of Technology, Japan)
 Shohei Shimizu (Osaka University, Japan)
 Peter Spirtes (CUJ, USA)*
 Milan Studeny (Institute of Information Theory and Automation, Czech Republic)*
 Joe Suzuki (Osaka University)**
 Maomi Ueno (University of Electro Communications)**
 Yi Wang (Institute of High Performance Computing, Singapore)
 Takashi Washio (Osaka University, Japan)*
 Changhe Yuan (Queens College/CUNY, USA)
 Jiji Zhang (Lingnan University, Hong Kong)
 Kun Zhang (University of Southern California, USA)

本授業の目的

- ベイジアンネットワークの推論、学習アルゴリズムについて学ぶとともに その下位モデルについても学ぶ。
- ベイジアンネットワークの世界最先端レベルまで学びたい！！
- ベイズ人工知能、機械学習、文書処理、数理情報科学、確率情報処理に興味ある諸君は履修すべし！！

内容

1. ガイダンス
2. ベイズ原理はだれが作ったのか？
3. 確率とベリーフ
4. グラフ理論
5. ベイジアンネットワーク
6. ジョイントツリーアルゴリズム
7. 近似推論
8. ベイジアンネットワークの学習
9. 大規模ベイジアンネットワークの学習

教科書(これで行う)



ベイジアンネットワーク
 植野真臣 著
 コロナ社

正誤表
<http://www.ai.is.uec.ac.jp/maomi/index.html>

成績の付け方

- 出席 欠席 3回以内 成績 Cが確保
- 普段の演習, レポート のできに応じて, S, A, Bの成績が割り振られます。
- 基本レジュメは <http://www.ai.is.uec.ac.jp/lecture/> に置いていきますので各自であらかじめ印刷して持ってきてください。