

## 多層Bayesian networkによる感性的事象のモデリング

|     |   |
|-----|---|
| 著者  | 福多 賢太郎, 小山 哲平, 出村 充, 魚住 超   |
| 雑誌名 | サテライト・ベンチャー・ビジネス・ラボラトリー年報   |
| 巻   | 6   |
| ページ | 45-46   |
| 発行年 | 2004  |
| URL | <a href="http://hdl.handle.net/10258/326">http://hdl.handle.net/10258/326</a> |

# 多層Bayesian networkによる感性的事象のモデリング

福多賢太郎、小山哲平、出村 充、魚住 超  
室蘭工業大学 情報工学科

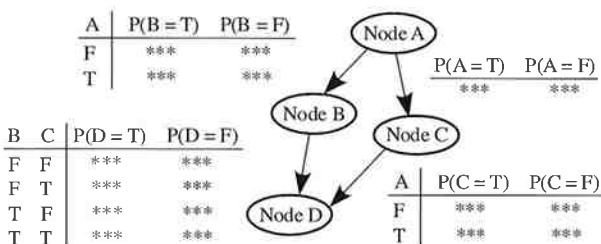
## 1. はじめに

我々は定量的及び定性的な側面から事象の因果関係をモデル化する Bayesian network(BN)<sup>[1]</sup>を活用することによって、現在または今後取り扱っていくべき問題領域に対して、新しい視点を提供できると考えている。特に広義には人の心理や生理活動を含み、具体的には嗜好、感情、行動といった感性的事象を一般的な枠組みの中で記述することは、人とコンピュータとの間のより親密なインタラクションの構築や、感性産業を推進するための一つの技術となる可能性がある。

しかし、BN をこれらの問題領域に適用するためには、計算量的な問題やモデル構築、推論における手続きなどに関して、解決しなければならない問題が多数存在する。ここでは、上記の問題に対する我々のアプローチを記述し、多層 Bayesian network を新たに提案する。多層 BN の具体例として、生理及び心理情報を包括的に取り扱った、VDTs(Visual Display Terminals)の使用に伴う視知覚疲労に焦点を当てた BN<sup>[2]</sup>を示す。

## 2. Bayesian network

BN は統計的データを構成する各項目を、グラフ論におけるノードに対応させて構成するグラフ構造の一種である。そして、定性的な因果関係を有向非循環グラフ(DAG : Directed acyclic graph)を用いて、ノード間に存在する確率的な性質を定量的な条件付確率分布を用いて視覚的に表現する。また、数学的には有限個の確率変数  $X(1), X(2), \dots, X(N)$  の同時確率分布  $P(X(1), X(2), \dots, X(N))$  をコンパクトに表現するための手法といえる[図 1]。



$$P(A, B, C, D) = P(A) P(B|A) P(C|A) P(D|A, B, C)$$

$$\Rightarrow P(A, B, C, D) = P(A) P(B|A) P(C|A) P(D|B, C)$$

図 1 Bayesian network と同時確率分布

BN の構築は、①ノードの作成(統計的データの離散化)、②ネットワークに組み込むノードの選定、③グラ

フ構造と確率分布の学習という、3 ステップに分けて考えることができる。

①、②では、異常値を取り除くといった前処理を施した上で、データの性質を反映するように数段階に離散化を行ってノードを作成し、ネットワークに組み込むノードを選定する。これはネットワークが有用となるかを左右する重要なプロセスであり、また、作成者のオリジナリティが最も反映される部分ともいえる。

③では、モデルの探索手法とモデルの良し悪しを判断する選択基準にバリエーションがある。一般的にモデル探索手法としては、K2 Algorithm, Greedy Search, Genetic Algorithm を活用したもの、全探索などが用いられ<sup>[3]</sup>、選択基準としては MDL, AIC, BIC, BDe といった情報量基準が広く用いられている<sup>[4,5]</sup>。我々は、モデル化の局面の違いによって、全探索と AIC や MDL の組み合わせを用いる場合と、Simulated Annealing (SA)<sup>[7]</sup>と BDe<sup>[5]</sup>との組み合わせを用いる場合の 2 つのアプローチをとる。

問題領域に対する知識が未知の場合は、探索手法に全探索を用い、選択基準はモデルが有すべき性質の違いによって使い分けることが妥当である<sup>[4]</sup>。しかし、DAG の数はノード数に応じて超指数的に増大するため<sup>[9]</sup>、ノードの持つ離散値や統計データ数によっても変化するが、一般的なコンピュータを用いる場合は 10 ノード前後の全探索までが現実的と思われる。

より多くのノード数を持つ BN の構築を望む場合は、ヒューリスティック探索を用いる必要があるが、探索ルールの与え方によって探索領域を著しく限定してしまうこともあり、BN の最適性の評価などが難しくなる。我々はこのような全探索が困難な場合の構造探索に、対数的アニーリングを用いた場合に最適性が保障される SA を用いる<sup>[7]</sup>。さらに SA を開始する際の初期構造に、問題領域に対する既存の知識(経験則や専門家の知識)を利用することを考える。そして、このような事前知識を効率よく探索に反映させるために BDe の性質に着目し、SA と BDe の併用に関して、その最適性や有効性などに関して研究を続けている<sup>[6]</sup>。

## 3. 多層 Bayesian network

2 節で挙げたような制約を越えて、より多くのノードを持つ BN を構築する方策を考える。一つは、ネットワークの主要部分を構成する 10 個前後のノード群に対して全探索を行い、残りのノードに関してはノード

の追加学習を行う方法である。しかし、ネットワークの基本構造を変更することが困難な点や、ノード数の増加に伴う確率推論に掛かるコストの増加、ネットワークの巨大化による視覚的に簡潔な理解が妨げられるといった問題が考えられる。

我々はこのような問題に対して BN を多層的に積み上げるといったアプローチを提案する[図 2]。まず統計データを複数のグループに分割する。このような手法としては、データの性質を吟味した上で任意に分割したり、ラフ集合<sup>[8]</sup>を用いたりする方法などが考えられる。そしてグループ毎に BN を構築し、各 BN に対して代表的なノードを算出し、これらノード群に対して再度 BN を構築する。このような形態をとることによって大局的な推論には上位の BN を使用し、あるノードの詳細な状態を知りたい場合は、そのノードに密接に関連したノード群で構成された下位の BN を用いて推論を行うという使い分けが可能となる。個々の BN は比較的小規模となるため視覚的理解はしやすく、また、推論のコストを下げることに繋がると思われる。

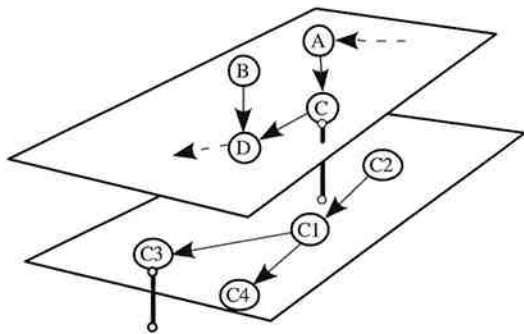


図 2 多層的 Bayesian network

このようなアプローチを理解するための具体例として、我々の行った VDT の使用に伴う視覚疲労問題を取り扱った BN<sup>[2]</sup>を示す。

#### 4. VDT 作業中の視覚状態の表現

図 3(上層)に示す BN は、部屋の照度や作業時間といった環境情報、VDT の解像度や輝度などの機器情報、瞬目や眼球運動といった生理情報、アンケートによる心理情報を包括的に扱った、VDT 作業中の視覚状態の因果関係の一部である<sup>[2]</sup>。3つの疲労状態(視覚疲労、身体的疲労、精神的疲労)はそれぞれ5つの項目を持つ主観的アンケートによって評価されており、図3(下層)は視覚疲労に焦点を置いた BN である。

大局的な視点で、あるノードの生起確率を求めたい場合は、より上層の BN に対して信念伝播法<sup>[1]</sup>に代表される推論アルゴリズムを用いる。これにより、例えば、部屋の照度という環境情報を得たときの、視覚

疲労の有無などを推察することが可能となる。そして、視覚疲労があると推論された時、下層にある BN を用いて推論を行うことによって、視覚疲労の詳細を推察することが可能となる。

現段階では2層間で共有されているノードをどのように表現し、上層から下層までの推論を円滑に行うための推論機構の実装に関して課題が残されている。しかしながら、ネットワークの構築から、階層的な視覚化までを一貫して行うシステムを構築することによって、多くの興味深い知見を得ることができると考えている。

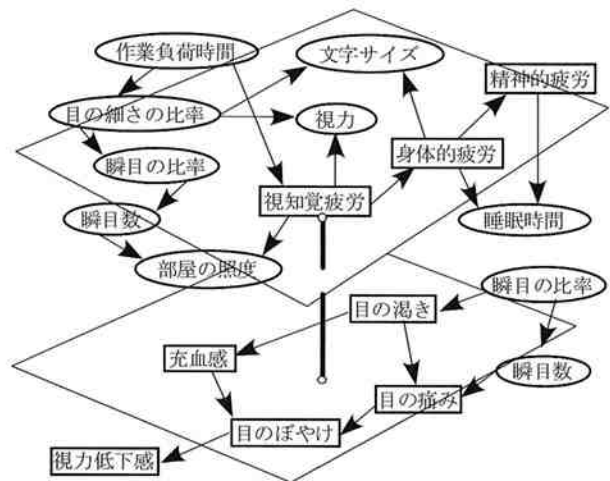


図 3 視覚疲労を中心とした多層的 Bayesian network

#### 参考文献

- [1] Pearl, J., Fusion, Propagation and Structuring in Belief Networks, Artificial Intelligence, Vol. 29, 1986
- [2] 小山哲平, 福多賢太郎, 魚住超, ページアンネットワークを用いた VDT 作業中の視覚疲労の表現, 第 37 回計測自動制御学会北海道支部学術講演会論文集, pp. 153/154, 2005
- [3] Russell, S. and Norvig, P., Artificial Intelligence: A Modern Approach, 2<sup>nd</sup> Edition, Pearson Education, Inc., 2002
- [4] 柴田里程, 情報量基準による統計的モデル選択, 電子情報通信学会論文誌, Vol. J83-A, No. 6, pp. 605/611, 2000
- [5] Heckerman, D., A tutorial on learning with Bayesian networks. Technical Report MSR-TR-95-06, Microsoft Research, Mar. 1995; revised Nov. 1996
- [6] 出村充, 福多賢太郎, 魚住超, BN における BDe を用いた探索手法の収束性に関する研究, 第 37 回計測自動制御学会北海道支部学術講演会論文集, pp. 159/160, 2005
- [7] Kirkpatrick, S., Gelett Jr. C. D., Vecchi, M. P. Optimization by Simulated Annealing, Science, 1983
- [8] 森典彦, 田中英夫, 井上勝雄, ラフ集合と感性, 海文堂出版株式会社, 2004
- [9] Harary, F. Palmer, E. M. Graphical Enumeration, New York: Academic Press, 1973.