

ベイジアンネットワーク

本村 陽一[†]

[†] 産業技術総合研究所 デジタルヒューマン研究センター

E-mail: [†y.motomura@aist.go.jp](mailto:y.motomura@aist.go.jp), <http://staff.aist.go.jp/y.motomura/>

あらまし ベイジアンネットワーク (以下ベイジアンネットと略す) は不確実性のもとでの予測や意思決定に用いられる確率モデルの一つである。このベイジアンネットを実際に用いるためには大量のデータからの統計的学習が重要になる。問題領域を表すベイジアンネットを構築した後、確率的な推論や予測を行うことができる。本稿では、ベイジアンネットに関するモデル、確率推論、統計的学習などの研究の流れを概説し、知的情報システムへの応用や応用システムの事例についても紹介する。

キーワード ベイジアンネット, 確率推論, 統計的学習, 障害診断, ユーザモデル

Bayesian networks

Yoichi MOTOMURA[†]

[†] National Institute of Advanced Industrial Science and Technology, DHRC

E-mail: [†y.motomura@aist.go.jp](mailto:y.motomura@aist.go.jp), <http://staff.aist.go.jp/y.motomura/>

Abstract Bayesian networks are probabilistic models that can be used for prediction and decision making under uncertainty. In order to use Bayesian networks for practical applications, learning (model construction) from actual data is important. In this paper, we show a model, probabilistic reasoning, statistical learning of Bayesian network. Applications of Bayesian network are also introduced.

Key words Bayesian network, Probabilistic reasoning, Statistical learning, Trouble shooting, user modeling

1. はじめに

情報技術の応用範囲が飛躍的に拡大している中で、これまで計算機にとっては比較的扱いにくかった不確実で非明示的な情報を取り扱うことが必要になってきている。例えばインターネットの発達により、ユーザの年齢、職業などの基本属性、嗜好性、これまでした買物や質問のような行動履歴などのデータなどが集積されているので、そこからある顧客層に人気のある情報や商品、あるいは逆に、ある対象に興味を持つユーザ層を見つけ出し、こうした情報を積極的に使った情報サービスも考えられている。また、知的情報処理システムが実際の問題領域において観測データから自律的に学習を行い、不確実な情報からでも確率的な推論や予測を行う仕組みも期待されている。例えば、複雑なシステムにおいて障害が発生した時の状況を全てデータベースに記録しておき、後で何か障害が発生した時に自動的に原因となる可能性の高い要因を推定することが考えられる。

こうした要請に応えるために、統計データに基づいてモデルを構築する統計的学習理論や、不確実性のもとでの確率推論の研究が重要になっている。とくに意思決定理論に基づいてシス

テムを制御したり有用な知識を表現するには、比較的複雑な構造を持った確率モデルが必要になる。このような確率モデルの一つに、変数間の依存関係や因果関係を有向グラフで表すベイジアンネットがある (ベイジアンネットに関する詳細については、英語の教科書としては [1] ~ [4] が、また日本語で読める解説としては [5] ~ [10] などがあるので、さらに興味のある方は適宜参照していただきたい)。これは不確実性をふくむ事象の予測や合理的な意志決定、障害診断などに利用することのできる確率モデル (グラフィカルモデル) の一種である。最近、このモデル上での確率推論アルゴリズムの進歩や、不確実性を含む様々な問題への応用、ソフトウェアの普及などにより、様々な分野で期待が高まっている。本稿では、ベイジアンネットに関するモデル、確率推論、統計的学習などの研究の流れを概観し、知的情報システムへの応用や応用システムの事例についても紹介する。

2. ベイジアンネット

ベイジアンネット (Bayesian network, Bayesnet, belief network) とは複数の確率変数の間の定性的な依存関係をグラフ構造によって表し、個々の変数の間の定量的な関係を条件付確率

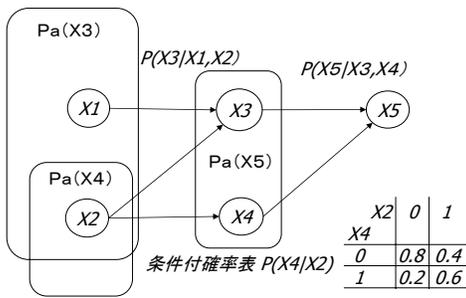


図1 ベイジアンネットワーク
Bayesian network

で表した確率モデルである。確率変数と、その間の依存関係を表すグラフ構造、条件付確率、の集合によって定義される。確率変数としては例えば「会議に關係するメールが来る」というような事象について定義し、その可能性に関して0から1の間の確率値をとる。「次の会議が行われる曜日」(月曜から日曜日)のように複数の状態と、それぞれの確率をとる場合もある。「前回の会議の曜日」のように観測の結果、状態を確定できる場合にはその状態の確率値は1、それ以外の状態については0となる。

変数はノードとして、変数間の依存関係は原因から結果となる変数の向きを持つ有向リンクで図示する。例えば、確率変数 X_i, X_j の間の条件付依存性をベイジアンネットワークでは $X_i \rightarrow X_j$ と表す。リンクの先に来るノード(この場合は X_j)を子ノード、リンクの元にあるノード(この場合は X_i)を親ノードと呼ぶ。

親ノードが複数あるとき子ノード X_j の親ノードの集合を $Pa(X_j) = \{X_1^j, \dots, X_i^j\}$ と書くことにする。 X_j と $Pa(X_j)$ の間の依存関係は次の条件付確率によって定量的に表される。

$$P(X_j | Pa(X_j)) \quad (1)$$

さらに n 個の確率変数 X_1, \dots, X_n のそれぞれを子ノードとして同様に考えると、全ての確率変数の同時確率分布は式(2)のように表せる。

$$P(X_1, \dots, X_n) = \prod_{j=1}^n P(X_j | Pa(X_j)). \quad (2)$$

こうして各子ノードとその親ノードの間にリンクを張って構成したベイジアンネットワーク(図1)によって、これらの変数間の確率的な依存関係がモデル化できる。先の例で言えば「前回の会議の曜日」と「会議に關係するメールが来る」という二つの事象と「次の会議のある曜日」の間に依存関係があるならば、この3つのノードを使ったベイジアンネットワークによって、その確率的な関係がモデル化でき、その上で予測などができるといふわけである。

変数が離散の場合、式(2)の右辺にある条件付確率は例えば $P(X_j = \text{Mon} | Pa(X_j) = \text{true}, \text{Mon}) = 0.8$ のような形で挙示することができ、子ノードと親ノードがとる全ての状態のそれぞれにおける確率値を定めた表(条件付確率表:CPT)を使う

ことで完全に表現できる。

一般的に書くと、親ノードがある状態 $Pa(X_j) = x$ (x は親ノード群の各値で構成したベクトル)のもとでの n 通りの離散状態 (y_1, \dots, y_n) を持つ変数 X_j の条件付確率分布は $p(X_j = y_1 | x), \dots, p(X_j = y_n | x)$ となる(ただし $\sum_{i=1}^n p(y_i | x) = 1.0$)。これを各行として、親ノードがとりうる全ての状態 $Pa(X_j) = x_1, \dots, x_m$ のそれぞれについて列を構成した表1の各項目に確率値を定めたものが X_j にとっての条件付確率表(CPT)である。

$p(y_1 Pa(X_j) = x_1)$	\dots	$p(y_1 Pa(X_j) = x_m)$
\vdots	\ddots	\vdots
$p(y_n Pa(X_j) = x_1)$	\dots	$p(y_n Pa(X_j) = x_m)$

表1 条件付確率表(CPT)
Conditional probability table(CPT)

このベイジアンネットワークのある一つの子ノードに注目した依存関係を、一つの目的変数(従属変数: Y)と、それに対する説明変数(独立変数: X)の間の依存関係として見ると、統計分野における回帰モデル、因子分析や共分散構造分析などの多変量解析、人工知能分野における決定木、ニューラルネットなどと比較して特徴を理解することができる(図2)。従来の多くの多変量解析的手法では、相関や主成分分析、因子分析のように変数間の線形の共変関係に基づいてモデル化が行われることが多い。グラフィカルモデリングへの拡張である共分散構造分析も目的変数にはガウス分布を仮定し、それに関する平均、分散パラメータを説明変数(の線形関数)によってモデル化する枠組みと理解することができる。データからの階層型ニューラルネットの学習は非線形関数(または写像)によるモデル化とみなすこともできる。

これに対してベイジアンネットワークは $X - Y$ 空間を条件付確率表にしたがって離散化し、個々の確率値を割り当てた不連続な確率分布によるモデル化である。その自由度は比較的高いものになっており、線形から非線形な依存関係まで柔軟に近似することができる。また各項目毎に十分な数の統計データがあれば、変数の各状態についての頻度を正規化して、各項目の確率値を求めることが容易にできる。決定木もベイジアンネットワーク同様に変数空間を分割するようにモデル化するが、分割の仕方がやや異なっている。

3. ベイジアンネットワークの確率推論

ベイジアンネットワークを使うことで、一部の変数を観測した時の他の変数についての確率分布を求めたり、確率値が最も大きい状態をその変数の予測結果として得ることができる。観測された変数の情報(e)から、求めたい確率変数(X)の確率値、すなわち事後確率 $P(X|e)$ を求め、それにより X の期待値や確信度最大の値(MAP値)、ある仮説の確信度(いくつかの変数が特定の値の組をとる同時確率)などを評価するわけである。先の例では「次の会議のある曜日」を予測したり、異常を観測したセンサーの状態からシステムの障害原因を推定するよ

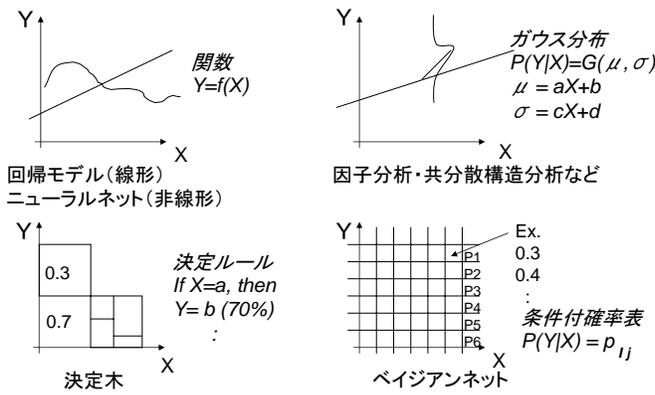


図2 ベイジアンネットワークと他の確率モデルの比較
Bayesian network and other probabilistic models

うな計算処理である。こうした確率計算に基づく推論が確率的推論と呼ばれている。

3.1 確率伝搬法

ベイジアンネットワークによる確率的推論は、i) 観測された変数の値 e をノードにセットする、ii) 親ノードも観測値も持たないノードに事前確率分布を与える、iii) 知りたい対象の変数 X の事後確率 $P(X|e)$ を得る、という手順で行なわれる。iii) における事後確率を求めるために、観測された情報からの確率伝搬(変数間の局所計算)によって各変数の確率分布を更新していく確率伝搬法 (belief propagation) と呼ばれる計算法が使われる。ここでは簡単に、図3の構造のもとでの計算の実行例を示す。

$X_1 \rightarrow X_2, X_2 \rightarrow X_3$ の間に依存性があり、条件付確率が与えられているとする。今、計算しようとしているノードを X_2 として、上流にある親ノードに与えられる観測情報を e^+ 、下流の子ノードに与えられる観測情報を e^- と書く。計算したい事後確率 $P(X_2|e)$ は、 e を e^+ と e^- に分け、 X_2 と e^- に注目してベイズの定理を使うと次のようになる。

$$P(X_2|e) = P(X_2|e^+, e^-) = \frac{P(e^-|X_2, e^+)P(X_2|e^+)}{P(e^-|e^+)}$$

また e^+ と e^- は X_2 を固定した時には条件付き独立になるので、 $\alpha = \frac{1}{P(e^-|e^+)}$ を X_2 の値によらない正規化定数とすれば、次のように変型できる。

$$P(X_2|e) = \alpha P(e^-|X_2)P(X_2|e^+) \tag{3}$$

このうち e^+ による X_2 への寄与分、つまり親ノードから伝搬する確率を $P(X_2|e^+) = \pi(X_2)$ と書く。これは、 $P(X_1|e^+)$ と X_2 のCPTを使って、 X_1 についての周辺化、次式(4)によって求めることができる。

$$\pi(X_2) = \sum_{X_1} P(X_2|X_1)P(X_1|e^+) \tag{4}$$

$P(X_1|e^+) = \pi(X_1)$ は観測値が与えられているならば、その値は決定できる。観測値がなく、さらに親ノードを持たない最上流のノードであるならば、事前確率を与える。その上流に親ノードを持つ場合には式(4)を再帰的に適用していけば最終的

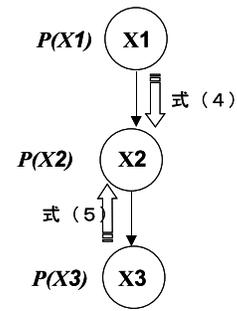


図3 簡単な構造での確率伝搬
Belief propagation on simple network

にはもっとも上流にあるノードによって、その値が求まる。

一方、式(3)の子ノード側の e^- の寄与分、つまり子ノードから伝搬する確率を $P(e^-|X_2) = \lambda(X_2)$ とすると、これを計算するためには、すでに定義されている条件付き確率 $P(X_3|X_2)$ を使って X_3 の全状態について周辺化する次式を用いればよい。

$$\lambda(X_2) = \sum_{X_3} P(e^-|X_2, X_3)P(X_3|X_2)$$

観測から得られる情報 e^- は X_2 の値によらず独立であることを利用すると、これは次式のように書き直せる。

$$\lambda(X_2) = \sum_{X_3} P(e^-|X_3)P(X_3|X_2) \tag{5}$$

ここで、 $P(X_3|X_2)$ は条件付確率表として与えられている。 $P(e^-|X_3) = \lambda(X_3)$ は観測情報が与えられているならば値が確定できる。また観測値がなく、その下流に子ノードを持たない下流のノードの場合には、無情報であるから一様確率分布であるとして X_3 の全ての状態について等しい値とする。また一般の構造のネットワークの場合、さらに下流に子ノードを持つならば、式(5)を再帰的に適用していけば、最終的にはもっとも下流にあるノードによって値が求まるので、やはり $\lambda(X)$ を計算することが可能である。

したがって、以上式(4),(5)を、式(3)に代入すればノード X_2 の事後確率が求まる。同様に次式によって任意のノードの事後確率も局所的に計算できる。

$$P(X_j|e) = \alpha \lambda(X_j) \pi(X_j)$$

ベイジアンネットワークのリンクの向きを考慮しないグラフ構造内の全てのパスがループを持たない時、そのベイジアンネットワークは singly connected なネットワークと呼ばれる。Singly connected なベイジアンネットワークの場合には、親ノード、子ノードが複数存在するような構造のネットワークでも、条件付独立性の性質を使うことで、各ノードについて上流からの伝搬、下流からの伝搬、上流への伝搬、下流への伝搬の4種について先の確率伝搬計算を図4のように行なうことで計算は完了する。またその計算量はノード数にたいして線形オーダーである[1]~[5]。

リンクの向きを考慮しないでネットワークを見たときに、どこか一つでもパスがループしている部分がある時、このベイジアンネットワークは multiply connected と呼ばれる。この場合には単

$$\Pr(X=x) = \alpha \lambda(x) \pi(x).$$

$$\pi(x) = \sum_u P(X=U=u) \prod_{U_i} \pi_{U_i X}(U_i),$$

$$\lambda(x) = \prod_{Y_j} \lambda_{Y_j X}(x),$$

$$\pi_{XY}(x) = \pi(x) \prod_{k \neq j} \lambda_{Y_k X}(x),$$

$$\lambda_{XU}(u) = \sum_x \lambda(x) \sum_{k \neq i} P(x|U) \prod_{k \neq i} \pi_{U_k X}(u_k).$$

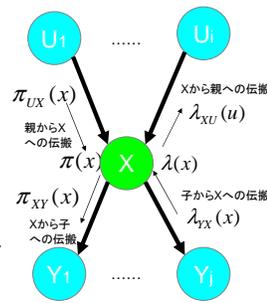


図 4 確率伝搬アルゴリズム
Belief propagation algorithm

純にリンクに沿って確率を伝搬してだけでは、その計算の収束性が保証できない。しかし 1990 年代のはじめに junction tree アルゴリズムと呼ばれる手法が開発されたことで、様々な構造にたいするベイジアンネットワークの有用性が高まった。この junction tree アルゴリズムを実装した Hugin (<http://www.hugin.com>) というソフトウェアも製品化されており、これを使ったシステムの実用化も進みつつある。

一方で multiply connected なグラフを変換せず、そのまま確率計算を行う統計学的な近似アルゴリズム [11] の理論研究が進み、中でも loopy belief propagation アルゴリズムと呼ばれるアルゴリズムが最近注目を集めている。

3.2 Junction Tree アルゴリズム

Multiply connected なネットワークの確率伝搬法を効率的に実行するための手法として junction tree アルゴリズムと呼ばれるものがある。

これはまず、適切な親ノードを併合する操作を繰り返してノードのクリークをクラスタとして生成し、元のベイジアンネットワークのノードをクリークとして結合した singly connected な木構造からなる junction tree (または moral graph) と呼ばれる無向グラフに変換する。次に、こうしてできた singly connected な木構造にしたがってクリーク毎に確率伝搬を行うことでやはり確率値が計算できる。複雑なネットワークの場合には、グラフ変換にかかる計算コストが大きくなるが、一度グラフ構造の変換に成功した後に何度も確率計算を行うような場合には非常に効率の良い確率伝搬を実行することができる。

ただし junction tree アルゴリズムは、ノード数が増え、グラフ構造が複雑になるにつれ、変換操作自体の計算コストが無視できない問題となる。とくにネットワーク構造が常に変わらない場合には、一度だけ変換を行えば良いが、状況によってネットワーク構造が変化する場合にはその都度グラフ構造を変換しなければならず、変換のための計算コストは深刻である。またグラフ構造の性質によっては効率のよい junction tree に変換できず、結果として巨大なクラスターが生じることも起こりえる。その場合にはクラスター内の確率計算のために多数の確率変数の全状態の組み合わせについての計算が必要のために計算量が増大する。

3.3 Loopy belief propagation

確率伝搬法を強引に multiply connected なベイジアンネット

トに適用する方法が loopy belief propagation と呼ばれている。Multiply connected なベイジアンネットワークに対して強引に局所的な確率伝搬法を繰り返し適用してみると、経験的に次のような良好な性質、グラフ構造の上の局所計算の相互作用によって部分の結果が徐々に全体に広がる場合には最終的にはほとんどのノードが正しい確率値に近い近似解に収束すること、また近似的な解に収束しないような場合の多くはノードの値が振動し、正しい解でないことも比較的容易に判別できること、などが実験的に示された [16]。

この loopy belief propagation はベイジアンネットワークを含むグラフィカルモデル一般について、アルゴリズムの改良やその収束性の解析などに関する理論研究 [12] ~ [15] が最近盛んであり、この中で統計力学や情報幾何学が重要な役割を果たしている。また日本でも若手研究者による研究が進められ、重要な国際会議の一つである Neural Information Processing Systems (NIPS) において日本から提案したワークショップが開催されるなど活発な活動が行われている。

4. ベイジアンネットワークの統計的学習

実際にベイジアンネットワークを応用しようとした時に最初にぶつかる壁はどのようなベイジアンネットワークを構築すれば良いかという問題であり、適切なモデルを構築する手法が本格的な実用化の鍵になっている。ここではベイジアンネットワークの学習の中心となる、条件付確率の学習と、グラフ構造の学習について述べる。

4.1 条件付き確率の学習

条件付確率表において、 X, Y のとりうる値の全ての組み合わせについてデータが存在する場合を完全データと呼び、そうでないものを不完全データと呼ぶ。完全データの場合にはデータの頻度により CPT の全ての項を埋めることができる。例えば簡単のため確率変数が真偽二値とすると、親ノード群 $\pi(X_j)$ がある値をとる時の X_j が真であった事例数を n_j^t 、偽であった事例数を n_j^f とする。 $\pi(X_j)$ を与えた時に X_j が真となる条件付確率が $P(X_j = 1 | \pi(X_j)) = \theta_j^*$ であったとするならば、このような観測結果 (n_j^t, n_j^f) が得られる確率は $\binom{n_j^t + n_j^f}{n_j^t} \theta_j^{*n_j^t} (1 - \theta_j^*)^{n_j^f}$ になる。条件付確率の学習では、この θ_j^* をデータの頻度 n_j^t, n_j^f から推定することになる。データの数が多く、 n_j^t, n_j^f が十分大きい場合には、最尤推定により $\tilde{\theta}_j = n_j^t / (n_j^t + n_j^f)$ とすることができる。

一方、全ての起こり得る組合せのデータを持たない不完全データの場合には、最尤推定量を求めることができない。とくに実際の問題に適用するにあたっては確率変数のとりうる値が多様であり、状態数の増加にともなって条件付確率表 (CPT) のサイズが増大する傾向がある。データから条件付確率を決定するためには、各項に十分な数のデータが必要になるが、CPT の全ての項目について一様に分布したデータが得られるとは限らない。とくにいくつかの項についてはデータが欠損している不完全データしか利用できない場合も多い。こうした不完全データの場合にはまず、未観測データについての確率分布を推定し、さらにその分布によって期待値計算を行なうことが必要になり、そのために EM アルゴリズムや MCMC 法が用いら

れる．しかし理論的には一見解決済みのようであるが，実際には繰り返し計算に時間がかかることや収束精度などの困難が多く，実用化は進んでいない．さらに実際の問題領域において収集したデータの分布には偏りや依存性があることなどからより現実的な解決策が望まれているのが現状である．そのための現実的なアプローチの一つとしては，Bound and Collapse 法やエントロピー最大化に基づく方法が提案されているが，欠損部を含む全予想サンプル数を事前に仮定する必要があるなど問題も残っている．また筆者は不完全データからの学習において，連続確率分布を表すニューラルネットワークをはじめにデータから学習し，それを用いて条件付確率表の欠損値を補完する手法とそのためのソフトウェアを開発している [19]．

4.2 グラフ構造の学習

ベイジアンネットワークの構造学習アルゴリズムとしては現実的な時間でグラフ構造を探索するためのヒューリスティクスを用いた K-2 アルゴリズム [17] があり，本システムもこれを基本にしている．子ノード 1 つを根，これに接続する親ノード群を葉とした木に注目すると，ベイジアンネットワークはこの木が複数組み合わさったものになっている．そして条件付確率分布はこの局所的な木のそれぞれについて一一定義される．そこでグラフ構造の決定は各子ノード毎に最適な局所木を探索する Greedy アルゴリズムとして実現できる．つまり，(a) 子ノードを定義，(b) 子ノード毎に候補となる親ノードの集合を与える，(c) 各子ノードごとに親ノードと条件付き確率を決定，(d) 情報量基準に基づき，最適な局所木を子ノード毎に探索する，という手順でベイジアンネットワークを構築するのが K-2 アルゴリズムの基本である．また (d) の手続きにおいては MDL の適用 [18] が一般的であるが，ベイジアンネットワークの構造学習の研究では，他にも様々な情報量基準の提案が行われている．

5. ベイジアンネットワークの応用

ベイジアンネットワークによる確率的推論の応用として成功している例としては複雑なシステムの障害診断がある．また最近ではユーザに適應するソフトウェアエージェントのためのユーザモデリングに利用する例なども増えてきている [8]．故障診断の例として Hewlett-Packard 社のプリンタ障害診断と Dynasty 社の汎用トラブルシューティング用ソフトウェアパッケージを，またユーザモデリングの例としてマイクロソフト社で研究されているプロトタイプソフトウェアを例にあげる．

5.1 障害診断への応用

ベイジアンネットワークのソフトウェアを開発している Hugin Expert 社と Hewlett-Packard 社のカスタマーサポート R&D は Systems for Automated Customer Support Operations(SACSO) プロジェクトという共同開発により，HP 社のプリンタに関する障害診断・発見システムを開発した [20]．他にも障害診断システムへのベイジアンネットワークの応用は NASA や Intel, Nokia 社をはじめ多くの例がある [22] が，この SACSO プロジェクトの場合は，プリンタのような民生品で，エンドユーザがアクセスするカスタマーサポートにおけるシステムとして実用化した点が目を引く．

5.2 ユーザモデリングへの応用

ベイジアンネットワークを用いたユーザモデリングの実証的なプロジェクトとして先駆的なものはマイクロソフトの Lumière Project [21] である．ベイジアンネットワークはある機能がユーザにとって適切である確からしさを推定するために用いられる．同様のアプリケーションとして Windows 上の代表的なソフトウェア Outlook をもじった LookOut という，1998 年頃からマイクロソフトリサーチの社内で試用されているプロトタイプを紹介する．LookOut は Windows のメールソフトである Outlook を使うユーザの操作をモニターし，ユーザが新しいメールを開くとそれを検出して，そのメールの内容を読み取り，スケジュール調整を行うためにカレンダーを開くかどうか，またいつ，どのようにユーザ支援動作を行うのが最も良いかなどを判断する．

LookOut がメールの内容を読み取る時の確率推論は次のようになる．それぞれのメールが到着すると LookOut はユーザが Outlook のカレンダーとスケジュールサブシステムを使う可能性を確率として計算する．この確率はメールヘッダの内容（メールの差し出し人や時間など）とテキスト文中に含まれるパターンによる確率推論で計算できる．この確率と，サービスを提供した場合にユーザがどれだけありがたいかという効用，動作のコストを考慮した期待効用が最適なアクション（場合によっては何もしないこと）などを決定する．

各アクションの期待効用の値によって，ユーザを支援する動作を行うか，カレンダーの表示とスケジュール入力を自動的に実行するかどうか，またはその必要がないので邪魔をしない，といったいくつかの候補の中から最良の動作を決定する．この推論の中で LookOut は人がミーティングなどを記述するとき用いる典型的な語句，“Friday afternoon”，“next week”，“lunch”などを検出する．それらの表現から想定される様々な場合の可能性を考慮して，スケジュールの決定に関連する可能性を確率として計算し，適切なユーザ支援動作を決定する．

もしも特定の日時と相手とのミーティングを行う確率が非常に高い場合には，その内容に基づいたサブジェクトと内容を入力済みの Outlook が起動し，この内容で良いか，あるいは修正するかをユーザに確認する．もし他のスケジュールとぶつかっていたら代替案として他の時間を探してからそれを表示する．一方，メールの情報が特定の日時を示している確率や，ミーティングを表す確率が低く，LookOut が起動するスケジュールリングサービスの期待効用が低い場合には，単にもっとも必要と思われる時期のカレンダーを表示するだけにとどめ，それ以上の余計な動作はしない．

ユーザが何をしたいと望んでいるかを確率と効用を通じてモデル化しているところ，そして観測だけからは確定できない事象に基づく不確実な意思決定を確率推論を用いることで実行しているところが，このシステムのポイントである．また，支援動作をキャンセルするなどのユーザからの適合性フィードバックを通じて，システムが学習できる点も非常に先進的であ

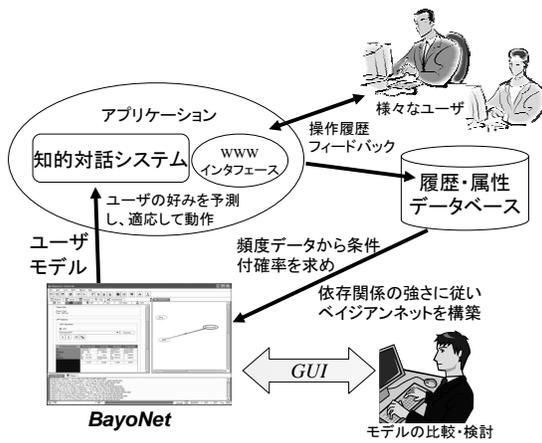


図5 ユーザの個性を自律的に学習する応用システム

る(注1)。こうした自律的な判断の信頼性が向上すれば、複雑な機能を提供するシステムの場合でもユーザは比較的単純な操作を行うことでも所望のサービスを受けられるようになると期待され、運転中のドライバーに対するアシストなどではとくに重要な技術となるだろう。

6. おわりに

計算機の高速化と、大量データが現実利用可能となってきたことで、情報システムが実際の問題領域における観測データから自律的に学習することも現実的になっている。身近な例では、送られて来たスパムメールを学習して自動分類を行うプログラムや、インタ-ネット検索システムである Google の WWW ページのランク付けなどにもベイズ確率に基づく確率的推論技術が利用されている。インターネットの発達により、ユーザの年齢、職業などの基本属性、嗜好性、などと、これまでした買物や質問のような行動履歴などのデータなどが大量に集積され、これらの因果関係を積極的に活用したサービス提供や問題解決、有用な知識の発見などを行うことも期待されている。これは従来から決定木などを用いたデータマイニングによる解析が行われている分野でもあるが、ベイジアンネットのように解析結果を知識モデルとして構築し、推論まで自動的に実行できるシステムに発展することはさらに大きな可能性を秘めている。これまでの情報処理システムはプログラマが設計した範囲でしか機能できなかったのであるが、統計的学習と確率的推論技術を用いることによってシステムの稼動中に得られた統計データにしたがいその動作を最適化し、様々な状況に適応しながら知的な処理が行える可能性が出てくるのである。日本でもベイジアンネットソフトウェア BayoNet(<http://www.msi.co.jp>) [19] が、そうした自律的に学習する知的情報処理システムを実現するために開発され、現在いくつかの研究プロジェクトでも使用されている(図5)。

今後は新しいセンサ技術や、ネットワーク環境、有機的に結合したデータベースなどの発展と連携した形で、ベイジアン

ネットが統計的学習に基づく有用な知識のモデル化の一つとして活用され、その上での確率的推論技術が計算機の新しい利用技術として確立されることが期待される。

文献

- [1] J. Pearl: "Probabilistic reasoning in intelligent systems", Morgan Kaufmann (1988).
- [2] F.Jensen:"An Introduction to Bayesian networks", University College London Press (1996).
- [3] E.Castillo,J. Gutierrez and A.Hadi:"Expert systems and probabilistic network models", Springer-Verlag (1997).
- [4] R.G. Cowell, A.P.Dawid, S.L. Lauritzen and D.J. Spiegelhalter:"Probabilistic Networks and Expert Systems", Springer-Verlag (1999).
- [5] S. Russell and P. Norvig: "Artificial intelligence, A modern approach", Prentice Hall 1995, 邦訳: 古川康一監訳, "エージェントアプローチ人工知能", 共立出版 (1997).
- [6] 本村陽一, 佐藤泰介: "ベイジアンネットワーク-不確定性のモデリング技術-", 人工知能誌, vol.15, no.4 (2000), 575.
- [7] 本村陽一: "ベイジアンネットワーク", 信学誌, vol.83, no.8 (2000), 645.
- [8] 本村陽一: "ベイジアンネットワークソフトウェア", 人工知能誌, vol.17, no.5 (2002), 559.
- [9] 本村陽一: "ベイジアンネットによる確率的推論技術", 計測と制御, vol.42, no.8 (2003).
- [10] "ベイジアンネットセミナー BN2002", <http://staff.aist.go.jp/y.motomura/bn2002/> (2002).
- [11] 樺島祥介, 田中利幸: 統計力学と情報処理, 電子情報通信学会誌, 83-8, 630/636, (2000).
- [12] Y.Kabashima and D.Saad: Belief Propagation vs. TAP for decoding corrupted messages, Europhys. Letter, 44-5, 668/674 (1999).
- [13] T.Tanaka: A theory of Mean Field Approximation, Advances in Neural Information Processing Systems, 11, 351/357 (2000), MIT Press.
- [14] Shiro Ikeda, Toshiyuki Tanaka, and Shun-ichi Amari: Information geometrical framework for analyzing belief propagation decoder, , Advances in Neural Information Processing Systems, 14, 407/414, (2002).
- [15] 田中和之: 確率推論に対する統計力学的アプローチ-クラスター変分法と信念伝搬アルゴリズム-, 計測自動制御学会 システム・情報部門学術講演会 SSL-2002 予稿集 (2002).
- [16] K. P. Murphy, Y. Weiss, M. I. Jordan: Loopy Belief Propagation for Approximate Inference: An Empirical Study, Proceedings of Uncertainty in AI (1999).
- [17] G. Cooper and E. Herskovits:"A Bayesian method for the induction of probabilistic networks from Data", *Machine Learning*, vol.9 (1992),309.
- [18] J. Suzuki:"A construction of Bayesian networks from databases based on an MDL principle", *Proc. of the 9th Conference on Uncertainty in Artificial Intelligence* (1993), 266.
- [19] 本村陽一: ベイジアンネットソフトウェア BayoNet, 計測と制御, vol.42, no.8, (2003).
- [20] F. Jensen,U. Kjarulff,B. Kristiansen, H. Langseth, C. Skaanning, J. Vomlel and M. Vomlelova : "The SACSO methodology for troubleshooting complex systems", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing (AIEDAM)*, vol.15 (2001),321.
- [21] E. Horvitz, J. Breese, D. Heckerman, D. Hovel and D. Rommelse: "The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users", in *14th National Conference on Uncertainty in Artificial Intelligence* (1998).
- [22] <http://excalibur.brc.uconn.edu/~baynet/fieldedSystems.html>

(注1): 従来の製品に内蔵されていたアシスタント機能が余計なお世話と感じていたユーザにとっても、こうしたより控え目な支援機能と適合性フィードバックによるユーザへの適応機能などは好ましいと感じられるのではないだろうか。