

大規模データからの日常生活行動予測モデリング

— 実サービスを通じたベイジアンネットワークの学習と推論 —

本村 陽一

ベイジアンネットワークの統計的学習、確率推論技術とユーザモデリング技術、大規模データ収集技術を要素技術として構成した生活行動予測モデルの構築技術について述べる。また因果的な構造をグラフィカルモデルであるベイジアンネットワークを状況や文脈も含んだ大規模データから構築するための必然から生まれた「実サービスを通じた調査・研究」の概念についても議論する。

キーワード: ベイジアンネットワーク、統計的学習、確率推論、ユーザモデル、行動分析、知識循環

Daily life behavior modeling from large scale data

– Statistical learning and probabilistic reasoning of Bayesian networks through real services –

Yoichi Motomura

Daily life behavior modeling is discussed. This modeling framework consists of statistical learning, probabilistic reasoning, user modeling, and large-scale data collecting technologies. Bayesian networks can represent causality relationship as graphical structures. Such models should include situations and contexts of daily life behavior through real services. In order to collect large-scale data connected with them, we have to provide real services supported by many users. This concept is named “Research as a service” and discussed in this paper.

Keywords: Bayesian network, statistical learning, probabilistic reasoning, user model, behavior analysis, knowledge circulation

1 はじめに

情報処理技術の適用領域は益々拡大している。それにともない日常生活を支援する情報サービスが望まれてきている。そのためには様々な状況において人が何を目的としているか、という日常生活行動を計算論的レベルで記述したモデルが必要となる。この計算論的モデルを利用することで、ユーザの行動からその背後にある要求や期待している結果を予測し、システムがそれを速やかに実現することで日常生活を支援する新たなサービス開発が可能となる。またそうした人間との協調動作を日常生活中でシステムが実行し続けることで、これまでの実験室環境では得られなかった大量で意味のあるデータが獲得できるようになる。この大規模なデータを用いて、絶えずモデルを更新しつつサービスを運用し続ける循環を産み出すことができる。

しかし、こうした日常生活中においては不確実な情報（例えば、真か偽に断定できないような予測や不完全な観測情報）のもとでの情報処理が本質的に重要となり、これまでのシステムの記述方式で中心的な役割を果たしていた決定論的なアプローチから非決定論的アプローチへのパラダイ

ムシフトが必要になる。非決定論的アプローチとは、あいまいで不確実な情報をできるだけそのまま取扱い、計算するアプローチである。対象となる変数を確率分布として計算し、推論を行う確率的推論もその一つである^[1]。この確率的推論は事後確率を最大化するパターン識別器などではナイーブベイズや隠れマルコフモデル (HMM) として自然に用いられてきたものでもある。さらに意思決定理論に基づいてシステムを制御し、有用な知識を表現し、複雑な処理を行うためには、多数の変数からなる高次元確率分布の計算が必要になる。変数の数が膨大になると高次元確率分布の計算は困難になるため、局所的には低次元の確率分布を用いて近似する他なく、そのために変数間の関係を規定したグラフ構造を導入する。このようなグラフ構造を持つ多次元確率分布モデルとしてベイジアンネットワーク（以下ベイジアンネット）^[2]がある。ベイジアンネットは多変数間の依存関係を条件付き確率とネットワーク構造により規定した一般的なモデルである。またベイジアンネットは大規模データからの統計的学習によってモデルを構築でき、これもまた不確実性に対処するためには重要な性質となる。

産業技術総合研究所 サービス工学研究センター / デジタルヒューマン研究センター 〒305-8568 つくば市梅園 1-1-1 中央第2 / 〒135-0064 江東区青海 2-41-6

Center for Service Research/Digital Human Research Center, AIST Tsukuba Central 2, Umezono 1-1-1, Tsukuba 305-8568, Japan/Aomi 2-41-6, koto-ku 135-0064, Japan E-mail: y.motomura@aist.go.jp

Received original manuscript September 24, 2008, Revisions received January 13, 2009, Accepted January 13, 2009

本稿ではまず、非決定論的アプローチと確率モデリングについて述べた後、ベイジアンネットとそれを用いた日常生活における人の行動予測モデルの構築技術と応用事例について述べるとともに、これらを実現する過程において必然として構成されることとなった「Research as a Service (サービスを通じた調査・研究)」について議論する。

2 非決定論的アプローチの選択

実世界の問題においては、直接観測することができない対象（確率変数）について、その状態（値）やその可能性（確率）を知りたい。人間を対象にした計算処理にも必然的にこうした不確実性が入り込む。システムが何らかのタスクを実行する場合に、システムの中ではそのタスクがモデル化され、計算操作の対象になっているとみなせる。つまりプログラムは対象とするタスクのモデルと計算操作をプログラム言語によってコード化したものと理解できる。さらにどのユーザに対しても全く同じように動作するのではなく、ユーザによって動作を変えるようなことを考えると、システムの中では、タスクのモデルとユーザのモデルの2つを実装することが必要になる。通常はタスク（プロセス）のモデルに関しては明確に記述できることが多いが、ユーザに関しては人間に関する不確実性を取り扱うために多くの場合で非決定論的な計算モデルが必要になる。ユーザの意図や要求のような人間系内部の潜在変数は現在の所、明示的にモデル化することが難しく、こうした要素はそもそも非決定論的な枠組みで記述せざるを得ない。また多様なユーザが様々な状況においてシステムを利用する際の、システムがとるべき最適な動作を全て事前に規定しておくこともまた難しい問題である。システムが提供する機能はシステム設計者があらかじめデザインすべきであるが、システムのユーザが何を要求していて、提供された情報やサービスについてどのように受け止めたのか、システムの動作は正しかったのか、ユーザの期待とは違ったものであったのか、などはシステムの実行時や実行した後でないとわからない。つまり真に目の前のユーザにとっての最適な動作設計を事前に確定することは難しい。したがって目の前にいるユーザの期待や要求通りにシステムを動作させるためには、単に非決定論的な枠組みを用いるだけでも不十分で、さらにユーザの反応を実行時に予測した上で、その反応や評価を最適化するように、ユーザのモデルを動的に構築できる枠組みも重要になる。これが人間に関する不確実性である。

また計算対象としての情報が大量に出現し、計算可能な量との間に大きな隔たりが生まれることでも不確実性への対処が必要になる。例えば我々はインターネットの普及により、有限ではあるが、直接取り扱うことが困難な大量の

データというものに直面している。ある一つの Web ページが全てのユーザに読まれた頻度というものは計測可能ではあるが、これを全ての Web ページについて数え挙げるといような決定論的な取り扱いには現実的ではない。こうした場合に、Web ページの間の遷移確率というものを考え、これを定常的な確率過程として非決定論的にモデル化することによって、Google の PageRank は計算されている^[3]。つまり元の Web ページやリンク・被リンク構造は決定論的に記述されており、それを扱うコンピュータもまた決定論的なものであるにも関わらず、決定論的枠組みではなく、非決定論的なモデルを用いることで、記述量やデータ数の爆発に対応できているのである。こうした実世界や大量データ、人間を含めた系の不確実性に対処することが、これからの社会問題を解決するための人工知能システムには強く求められ、そこでは非決定論的なモデルとして問題を記述することが一つの解決策である。

不確実性を含む問題を非決定論的な計算モデルで記述したとしても、それを現在の決定論的なコンピュータで取り扱うということは、計算対象が何であるか、という計算論レベルのプロセスを、計算方法がどのように書かれ（アルゴリズムレベル）、どのように実行されるか（インプリメントレベル）、というレベルとは独立に考えるべきであるという Marr の計算論^[4]を想起させる。つまり、決定論的なシリコンチップのコンピュータの上で、決定論的なコンピュータ言語により記述されたプログラムで実行されているとしても、また先の Web の例のように元のデータやメカニズムが決定論的なものであったとしても、その計算対象のモデルとして非決定論的に考えることが有益な場面があるということである。トイ・プロブレムを対象とする限りは計算対象も決定論的に考えていても十分であるが、我々の目の前にある実問題を計算論的にモデル化しようとする、そこに内在する不確実性に対処するために非決定論的な枠組みで記述せざるを得ない。

3 ベイジアンネット

3.1 確率モデリング

非決定論的アプローチの一つとして、確率を用いる方法がある。確率を用いることで事象の不確実性を定量的にモデル化し、公理的確率論により厳密に取り扱うことが可能となる。観測可能な事象の確率値そのものは大量の観測データから得ることができ、観測不可能な事象については、ベイズ的な確率推論（ベイズ推定）によって推定することができる。これは条件付き確率によって、変数の不確実性と変数間の関係性をモデル化し、ある変数に関する不確実性を他の変数の情報から求めるものとして考えること

かりやすい。古典的なベイズ推定ではこの未知 (unknown) の確率分布を主観的な事前分布として扱うために、非ベイズの統計学者からは批判を浴びていたが、最近では大量データが取り扱い可能になったことで、この確率分布を大量の統計データから経験的に構成することが可能になり、多くの不確実性を持つドメインにおける実用的な方法として有望視されている。

例えば、完全に観測できない事象を扱う確率的な枠組について考えてみる。実世界には将来の天気や雑音混じりの信号、ユーザの意図のように確定値を得ることが難しい不確実な情報が多く存在する。これらを体系的に取り扱うために確率的な枠組を導入する。複雑な要因やノイズの影響などによって不確実性を含む対象を確率変数として X で表し、その変数がとりうる具体値を x_1, x_2, \dots, x_n と表すことにする。

次に変数間の依存関係を考える。例えば変数 X_i が x という値を取るならば、 X_j は y となる、という関係が成立しているとき、 X_j が X_i に依存していると考える (if $X_i = x$ then $X_j = y$)。現実には起きている複雑な事象を考えると、複数の変数間の依存関係は複雑になり、「if $X_i = x_1, \dots, X_i = x_n, \dots$, then $X_j = y$ 」のように明示的に全ての関係を列挙することはあまり現実的でない。また、たとえこのような if-then ルールを膨大に挙げたとしても実際には例外などがあり、必ずしも完全に状況を記述することは難しいだろう。そこで厳密な表現をあきらめ、主要な変数のみに注目し、ルールが成立する確信の度合いを定量的に表すために「 $X_i = x_i$ であるとき $X_j = y$ である確率は $P(X_j = y | X_i = x_i)$ 」という確率的な表現を導入する。二つの量 x, y の間の一意的な依存関係は、例えば関数 $y = f(x)$ によって表せるが、これと同様に、確率変数 X_i, X_j の依存関係は条件付確率分布 $P(X_j | X_i)$ によって表すことができる。これは X_i のとる値に応じて、 X_j の分布が影響をうけ、その依存関係の定量的関係が条件付確率分布 $P(X_j | X_i)$ で定められることを示している。

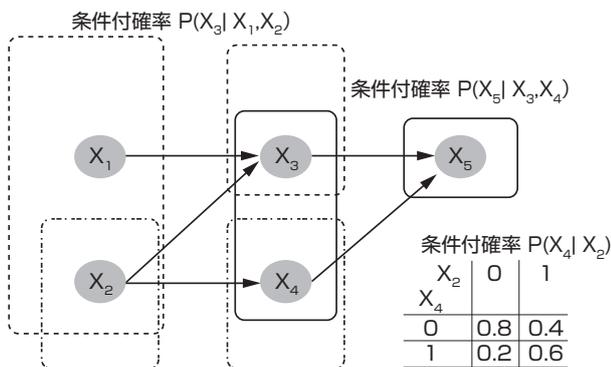


図1 ベイジアンネットワーク

さらに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、個々の変数間の定量的な関係を先の条件付確率で表したモデルがベイジアンネットである。説明変数、目的変数の区別なく、任意の変数の確率分布が効率よく計算できるのがベイジアンネットの特長でもあり、モデルは様々な用途に再利用することができる。

望ましい入力と出力の組からなるデータを与えることで、モデルやシステムの振る舞いを決定する枠組みが機械学習や統計的学習と呼ばれる。ベイジアンネットを実データからの統計的学習により構築することもできる。ベイジアンネットの上で行われる確率分布の計算は確率推論と呼ばれる。以降ではモデル、データからのモデル構築、確率推論のそれぞれについて簡単に述べる。

3.2 ベイジアンネットモデル

ベイジアンネットは数理的には確率変数をノードとするグラフ構造と、各ノードに割り当てられた条件付き確率分布群によってモデルが定義される (図1)。

各変数の条件付き確率分布は、離散的な確率変数の場合は条件付き確率表 (conditional probability table; CPT) として表現できる。このように条件付き確率を表として与えることで、確率分布を密度関数とパラメータで与えるよりも表現の自由度は高くなる。つまり、対象がどのようなものであるかが事前にはわからない対象に対する非決定論的なモデル化手法として有用である。

条件付確率が与えられる側の変数を子ノードと呼び、親ノードから子ノードの向きへ有向リンクを張る。このように変数とグラフ構造、条件付確率表により定義した非循環有向グラフをベイジアンネットモデルとして構築する。

3.3 データからのモデル構築

ベイジアンネットのモデルが大きなものになってくると、ネットワークの構造や全ての条件付確率表を手手で全て決定することはなかなか容易ではない。そこで大量のデータからの統計的学習によってモデルを構築する方法が必要となる。

学習に用いるデータセットが条件付確率表の全ての項目に対応する事例を含んでいる場合は完全データと呼ばれ、この場合には統計データを数え上げて頻度を得て、それを正規化したものが条件付確率値の最尤推定値となる。欠損がある不完全データの場合には各種の補完を行うことで条件付確率値を推定する。モデルのネットワーク構造も

$P(y_1 Pa(X_1) = x_1)$...	$P(y_1 Pa(X_1) = x_m)$
:	...	:
$P(y_n Pa(X_1) = x_1)$...	$P(y_n Pa(X_1) = x_m)$

表1 条件付き確率表 (CPT)

データから決定したいことがある。構造の学習はグラフ構造をある初期状態から探索するものになる。グラフ構造の良さをはかる評価規準としては、尤度の他に AIC や BIC、MDL などの情報量規準が用いられる。グラフのノード数が大きくなると探索空間は爆発的に増大し、グラフ構造を全て探索することは計算量の点から困難になるため、欲張り法 (Greedy algorithm) や各種のヒューリスティックを使い準最適な構造を探索することが必要となる。こうしたグラフ構造の学習アルゴリズムとして K-2 アルゴリズム^[5]がある。これは (i) 各ノードについて親ノードになりえる候補を限定しておく、(ii) ある子ノードを一つ選び、候補となる親ノードを一つずつ加えてグラフを作る、(iii) そのグラフのもとでパラメータを決定し、評価する、(iv) 評価が高くなった時だけ親ノードとして採用し、(v) 親ノードとして加える候補がなくなるか、加えても評価が高くならなくなったら他の子ノードへ移る、(vi) 全ての子ノードについて (i) - (v) を繰り返す、という探索アルゴリズムである。一般的には親の探索空間は組み合わせ的に大きくなるので、始めにノードを順序づけして候補となる親ノードの組合せを限定して計算量の増大を避ける工夫が必要になる。またグラフの探索部分(ii)、(v) とモデルの評価部分 (iii) をそれぞれ独立に考えることで様々な学習方法が考えられる。

ベイジアンネットを用いることで大量のデータから非決定的なモデルを統計的学習によって構築するアプローチの有効性が期待できる。しかし因果的構造を統計データだけから求めることは本質的に困難であり、またグラフ構造の探索問題は NP 困難である。そこで実際には変数候補や探索範囲の限定などを巧妙に行うことや適切な潜在変数の導入も必要である。

3.4 確率推論

グラフ構造を持つモデルは他にもあるが、その多くはデータを説明するグラフ構造を可視化するために用いられることが多い。一方、ベイジアンネットは離散確率変数と条件付確率表で構成されていることにより、モデルの中の任意の確率変数の確率分布推定を行う確率推論のアルゴリズムを非常に効率良く実行することができる。これが他のグラフィカルモデルにはない大きな特長であり、知的な学習システムを現実的な計算量で動作させるために重要な性質である。

ベイジアンネットの上の確率的推論は、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ておく、iii) 知りたい対象の変数 (X) の事後確率分布 $P(X|e)$ を計算する、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、変数間の依存関係にしたがって各変数の確率分布を更新していく確率伝搬法が用いられる。

ベイジアンネット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は singly connected なネットワークと呼ばれる。この場合には、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 種についての確率伝搬計算を行なうことで計算は完了する (図 2)。

この計算量はネットワークのサイズ (リンク数) に対して線形オーダーで済み、計算効率は非常に高い。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は multiply connected と呼ばれる。この場合には厳密解となる保証はないが、近似解法として確率伝搬法を適用することができ、LoopyBP 法と呼ばれている。

4 ユーザのモデリング

情報システムとユーザが対話的に処理を進めるということは、その情報システムは部分であって、動作主体としてのシステム全体としては、情報システムとそのユーザ、さらにそれらを取りまく環境や状況まで含めて考えなければならない。したがって制御対象としてのシステム全体をみると人間の行動や反応も計算対象の一部として考えるべきである。そこでユーザがある状況下では何を要求しており、またシステムの出力結果を得てどのように反応するのか、などを評価したい。そこでこうしたユーザの認知的な状態をシステム内で計算可能なモデル、ユーザモデルとして記述し、取り扱うことが必要となる。

機械学習の発展により機械 (プログラム) がデータにより学習する、つまりモデルはデータによって構築され、逐次的に修正されるというアプローチが可能となった。機械学習におけるモデル構築は、統計的検定を情報量規準による自動的なモデル選択として繰り返し実行し、最適なモデルを結果としている。すなわち、統計的に有意なモデルを広範な探索空間の中から機械学習により選んでいるわけで

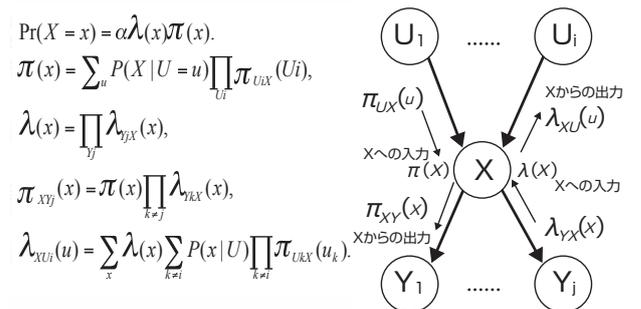


図2 確率伝搬法

$$\begin{aligned} \Pr(X = x) &= \alpha \lambda(x) \pi(x). \\ \pi(x) &= \sum_u P(X = u) \prod_{U_i} \pi_{U_i X}(U_i), \\ \lambda(x) &= \prod_{Y_j} \lambda_{Y_j X}(x), \\ \pi_{X Y_j}(x) &= \pi(x) \prod_{k \neq j} \lambda_{Y_k X}(x), \\ \lambda_{X U_i}(u) &= \sum_x \lambda(x) \sum_{k \neq i} P(x | U) \prod_{k \neq i} \pi_{U_i X}(u_k). \end{aligned}$$

ある。

さらに確率モデルを用いることによって、従来の心理学が普遍的な人間のモデルを扱おうとするために捨象していた分散としての要素、例えば個々に存在する個人の特性なども含めてモデル化することも可能になる。これは最近の人間中心設計やユーザビリティに配慮した情報処理において必要とされる個人適応、パーソナライゼーションを実現するためにも重要な観点である。確率モデルとしてベイジアンネットワークを用いた様々なユーザのモデル化が行われている^[6]。とくに人間の認知・評価構造をベイジアンネットワークとしてモデル化するために、臨床心理学やマーケティングで用いられているインタビュー法が適用されている^[7]。これによりシステムやサービスのユーザをモデル化し、確率推論を実行することで、嗜好性や意図を推定することが可能になる。具体的な例を7節で紹介する。

5 日常生活行動のモデリング

不確実性に関するコンピューティングの実際の例として、前節ではユーザのモデル化という観点で見えてきたが、様々な実サービスとして日常生活支援^[8]に目を向けてみると、ユーザとしての日常生活者のモデル化が重要になる。これまでにセンサを家の中に埋め込み、日常生活行動を分析するためにセンサハウスの研究開発も進められてきている^[9]。これまで計測されたデータの中にあるパターンを定常分布としてモデル化することで、外れ値を検出することで異常を判定するような応用はいくつか提案されているが、さらに応用を広げるためには、定常分布のモデル化だけでは不十分であり、ユーザの意図に応じて、効用や価値の最大化などを考える必要がある。つまり、直接観測できないユーザの意図や価値感や評価を、観測可能な行動から予測するための高次の推論が必要になる。そのためには、ある状況と行動に応じて結果がどうなるかという依存関係、因果関係をモデル化することが必要で、そのために行動のみならず原因となる変数も含めた包括的な観測データを収集し、得られた大量の変数間の関係から因果構造を探索することが必要である。

これはセンサ技術とモデリング技術が切り開く新たな行動分析学と見なせる。行動分析学^[10]は1900年代半ばにスキナーにより心理学における行動科学アプローチの一つの研究分野として確立され、その後は教育や臨床の場などで多大な貢献を示し、例えば障害児に対する高い教育効果を上げていることなどが知られている。そこでは人間の行動は先行条件と行動随伴性とよぶ、行動の結果として期待される環境の変化との三項関係より規定されるものと考えられる。そしてある特定の行動に注目した時の先行条件と行動

随伴性の間の因果関係を明らかにし、これを明示的にモデル化すること、その上で行動随伴性や先行条件を変化させることで行動の制御（行動変容）を実現するものである。

行動の因果を発見するために行動を観測したビデオ映像を解釈しラベル付けするような手段が必要となるが、これを人手で行う場合には手間と時間が膨大にかかることから、日常生活環境における自然な行動を効率良く分析することが難しい。また人手による解釈では、行動の制御変数として少数のものしか分析対象にできないため、日常生活行動の分析のためには自動的に大量の観測データを取り扱う技術が必要である。

そこで環境に埋め込んだセンサネットワークにより行動を自動的に観測し、これと統計的学習手法を活用することが考えられる。収集した大量のセンサデータからの統計的学習によりベイジアンネットワークモデルを構築することによって、行動随伴性の候補となる行動の理由・目的や、先行条件となる環境、状況の中での必然性などと結びつけることが可能になる。このようにして、日常生活行動を包括的に観測できるセンサ技術とそこで観測される大量データの中から因果的関係の強い変数を抽出するモデル構築技術の貢献が行動分析学を大きく発展させるという期待がある。これまでにベイジアンネットワークと超音波センサネットワークを用いた行動モデリングによる日常生活行動分析^[11]や、子供の傷害予防への応用^{[12][13]}などの研究が進められている。以下では子供の行動推定の例^[14]を紹介する。

部屋の中の人や物体に超音波発信機をつけることで超音波受信機を埋め込んだセンサルーム内の人や物体の各時刻における位置情報を x , y , z の座標データとして取得できる。また同時に部屋の天井部分に設置した魚眼カメラにより、部屋の中で人が行動する様子を動画として撮影する。この撮影された部屋の中の人や物体の行動に対する動画像を1秒ごとに人手でラベル付けを行う。例えば対象となる人が歩いている、座っている、立っているといった行動ラベルが部分的に付与されたデータベースを収集した。このデータを利用して、日常生活行動のモデル化と、それを用いた画像からの行動推定実験を行う。

行動をセンサや画像によってシステムが観測するものとして問題を考えると、これは一種のパターン識別の問題として定式化することもできる。実世界の日常において生成されるデータは人間の生活行動や生活環境を背景にしているから、データが発生する状態空間や頻度の偏りなどの性質は当然人間にとって解釈される意味が強く反映したものになっている。このようなデータが生成される空間に特有な制約や発生頻度の偏りを確率分布として扱うことができる。物理法則のようにその世界で成り立っている因果

構造を全て列挙することは記述量の点で困難であるが、その中の重要なものを確率として表現することは近似的に有効な手段である。そこで、実空間におけるこうした確率的な構造をベイジアンネットによりモデル化し、これによるベイズ推定に利用することが考えられる^[15]。

ベイズ推定では、複数のクラスラベルを C_i とし、信号パターン x に対する尤度 $P(x|C_i)$ と事前分布 $P(C_i)$ の両者を組み合わせた事後確率、

$$P(C_i|x) = P(x|C_i) P(C_i) / \sum_j P(x|C_j) P(C_j) \quad (1)$$

を最大化するクラスラベル C_i を決定する。これはベイズ誤り確率を最小にする最適な識別を可能にすることが知られている。データへの当てはまり具合は尤度で表し、事前知識は事前確率分布によって表される。そしてこの両者の積である事後確率を最大にするものを推定結果とすることで、データからの学習と事前知識が自然に統合されている。

クラスラベルの発生頻度が観測時間や観測場所に依存しているような場合、事前分布 $P(C_i)$ は状況 S に依存したものになっている。そこでこれを条件付確率 $P(C_i|S)$ として考え、これを式 (1) の $P(C_i)$ と置き換えて式 (2) の事後確率を最大とするクラスを識別結果とする。

$$P(C_i|x, S) = P(x|C_i) P(C_i|S) / \sum_j P(x|C_j) P(C_j|S) \quad (2)$$

式 (2) 右辺分子の第2項 $P(C_i|S)$ は、ラベル空間における状況 S の下での行動ラベル C_i の事前確率である。ここで、ラベル空間における確率的因果構造を考えることにする。場所や行動の系列の間の因果関係をベイジアンネットとして構築すると、例えば「状況 S で時刻 t に C_i^t という行動が起きたら、次の時刻 $t+1$ に C_i^{t+1} という行動が起きやすい」といった因果構造の形で事前知識を導入し、人が領域 S に入った時の行動の確率を $P(C_i^{t+1} | C_i^t, S)$ として表し、ベイジアンネットでモデル化することができる。実際にリビングルームを模した実験環境で子供が遊んでいる際の行動を観測したデータセットに対する統計的学習によってモデルを構築したところ、過去の行動の他に室内のソファや壁などの相対距離、移動速度などの依存関係が確認された。さらにある子供の行動データでベイジアンネットとナイーブベイズを学習し、これらを用いた式 (2) によるベイズ推定による別の子供の行動を推定したところ、ナイーブベイズのみの最尤推定では約50%未満の識別率であったものが、ベイジアンネットを用いたベイズ推定によって約60%~80%まで向上できることを示した^[14]。この行動推定アルゴリズムにより、日常生活行

動の観測画像などから行動ラベル付きデータを効率的に生成することが可能となった。

6 Research as a Service (RaaS)

日常における大規模データが観測できるようになったことで、統計的学習により複雑な問題に取り組めるようになった。しかし、統計的学習特有の問題として、モデルが高度で複雑なものになるにつれ、学習のために必要なデータ量が増えることがある。表層的に観測可能なセンサデータなどは比較的容易に取得できるが、人間行動の内部的状態は心理的なものであるため、被験者を用いたアンケート調査も必須になりコストが大きい。またデータを取得する上で、プライバシーの問題や、単に研究目的のためには協力が得られにくいという現実的な問題もある。またたとえ外部的要因で観測容易な事象だとしても、実際に使う場面において、状況依存性の高い説明変数を網羅的に収集するためには、データを観測する環境が日常的な利用環境とできるだけ合致するように統制しておく必要がある。

そこで、筆者はこうした問題に対して実サービスと調査・研究を一体化すべきであるとする「サービスとしての調査・研究 (Research as a service)」という概念を提唱している^[22]。これによって、行動分析学における人間の行動随伴性としての手段目的連鎖を明らかにし、状況依存性も含めて包括的にモデル化することが容易になる。そのためには、調査・モデル化の段階とそのモデルを用いた応用を切り離すことなく、情報サービスを社会の中で実行しながら、そこで得られる観測や評価アンケート、利用者のフィードバック（心理的調査）の結果を網羅的に収集する。これは古くはサイバネティクス、また信頼性工学ではデミングサイクルとして知られるPDCA (Plan, Do, Check, Action) サイクルを実問題を通じて回し続けることで、モデルを常に修正していくというものである。不確実性に対する本質的な解決のためには対象を実データによりモデル化し、そのモデルを用いて制御しながらさらにデータを収集する、というサイクルを永続的に続けるアプローチが必要になる。これは単に実データの収集だけにとどまらず、研究という視点で実フィールドに没入することで新しい価値・評価を生み出すという新しい研究のあり方^[16]に通じるものである。

そのためにも実サービスとして耐えられる製品として社会の中にインフラとして組み込める応用システムの実現が重要である。

7 ベイジアンネットの応用システム

確率推論アルゴリズムやモデル構築のアルゴリズムをコンピュータプログラムとして実装することでベイジアンネット

の応用システムが開発できる。筆者は2001年までのリアルワールドコンピューティングプロジェクト、IPA未踏ソフトウェアプロジェクトなどを通じて、2002年にベイジアンネットを大量データから探索し、その上で確率推論を実行することのできるソフトウェア BayoNet を開発した^{[17][18]}。このソフトウェアは民間企業へのライセンス供与、製品化もされたが、これを特定の問題解決に適用するためには高度な専門知識が必要なこと、利用手順が自明ではないことから、ソフトウェアを使いこなせるユーザがなかなか育たなかった。通常特定の目的のために開発されたソフトウェアであれば、ありえないことであるが、純粹に基礎的な数理モデル研究として生まれたベイジアンネットを実装したソフトウェアは、非常に幅広い目的のために適用することが可能であり、実用化できた時点で、あらためて、より価値の高い目的とのマッチングを検討するということが起こりえる。こうした状況の中で、ベンチャー開発戦略センターのタスクフォースが2003年に開始され、研究者自らがこの技術を使ったビジネスモデルの探索をはじめるといった機会を得た。この時点でアルゴリズムの洗練や高速化、推論精度の向上など要素技術としての研究課題も多く残っていたが、アウトカムが明らかでない状態で技術の先鋭化を進めることに抵抗を感じた。そこで、その時点での性能で十分対応可能な問題解決としてのアウトカムの探索を優先することにした。

ベイジアンネットを用いるメリットは、確率推論を行うことで、任意の変数に関する確率分布を求め、さまざまな条件における定量的な評価ができる点である。従来の多くの多変量解析的手法では、定量的な関係は、変数間の線形（線形独立）の共変関係に基づいてモデル化が行われることが多い。ベイジアンネットモデルでは定量的関係を条件付き確率表によって表わす。条件付き確率表では確率分布族を仮定することがなく、非線形、非正規な関係や交互作用も表現できる自由度の高いモデルになっている。また説明変数と目的変数を明示的に区別しないので、潜在変数の導入も容易である。つまり観測が得られない変数であっても、それを潜在変数として扱うことができる。これによりユーザや顧客の統計データを分析する際に、カテゴリとなる潜在変数を導入して、同じ行動をとるような集団の属性を抽出して顧客層を分類することができ、顧客セグメンテーションなどにも利用できる。

こうした性質はユーザや顧客の行動（Webブラウジング履歴など）や属性、状況に応じて、嗜好性にあった情報や商品を推奨するような応用にとっては非常に重要である。顧客やユーザにとって望ましいと思われる情報や商品を携帯電話やカーナビなどで表示する場合に、協調フィルタリングでは

状況依存性が反映できない。こうした動的に変化する環境での情報推薦技術は実空間で多様な状況変化が想定されるユビキタスコンピューティングにおいても重要である。

7.1 カーナビによるユーザ・状況適応型情報推奨

車を運転している途中で、どこかに立ち寄りたくなることがある。例えばある目的でドライブ中に、食事のためレストランに立ち寄ることを考える。これまでのカーナビではカテゴリを指定し、該当する全レストランが距離の近い順にリストアップされる。ユーザはリストの中から適切なレストランを見つけなければならないが、詳細な情報はタッチスイッチやリモコンを操作しないと確認できないためドライバーにとって望ましいレストランを見つけることは容易ではない。

そこでカーナビシステムがドライバーの嗜好性を表すベイジアンネットを利用し、これを使った確率推論によって、システムが運転中のドライバーに代わって自動的に適切な立ち寄り先を選定することが実現できれば非常に実用的な機能となる。人の嗜好は個性が大きく、また運転中の状況にも強く依存している。運転中には刻々と変化する状況の中で、その時々での最適な選択が必要である。

こうした状況依存性や個人差を表すために、変数間の複雑な依存関係と不確実性をモデル化できるベイジアンネットが有効に適用できる。そこで我々はユーザに適応してコンテンツを推薦するカーナビシステムの試作を行い評価した^[6]。このシステムは、ユーザの嗜好モデルをベイジアンネットとして車載情報システム内に持ち、レストランや音楽などコンテンツプロバイダより提供されるコンテンツがその時の状況、ユーザにどれだけ適切であるかを示すスコアを状況とユーザ属性を与えた時の条件付確率として計算し、このスコアの高い順に上位のコンテンツに限って提示するものである。実際に品川周辺の182のレストランに対し、6つの状況（シナリオ）の場合に行きたい店を選択させる質問を300名の被検者に対してアンケート実施し、収集したデータからモデルを構築した。品川周辺の182のレストランに対し、6つの状況（シナリオ）の場合に行きたい店を選択させた。選択手順は、最初に好きなカテゴリを質問し、そのカテゴリに該当する店を表示し、気に入らなければ次のジャンルを選ぶ、という現在の既存のカーナビと同様の選択方法をとった。選択レストランは複数回答であり、結果的に計3778レコードを得た。状況の属性数は12、レストラン属性は17、ユーザの属性数は12である。その結果として、図3のモデルを構築した。ユーザを表す属性ノードは4個、状況を表すノードは3個、レストランを表す属性ノードは6個の計13個の確率変数からなるモデルとなり、ある状況における特定のユーザが好むレストラン属性の確率分布が確率推論により計算できる。

図3のモデルでは運転歴が浅いドライバーではファミリーレストランやファーストフードチェーンのようなフランチャイズレストランが選ばれる確率が高く、逆に運転歴の長いドライバーはこうしたレストランを選ぶ確率は低い。これはフランチャイズレストランでは駐車場が整備されていることが多く、若者や初心者ドライバーに好まれる傾向を表している。さらに「運転歴」に加えて「次の予定（がある）」との交互作用があり、運転歴が長い場合でも、次の予定があり急いでいる状況ではフランチャイズレストランを利用する確率が上がるという傾向が反映されている。他にも予算レベルと車種の関係など直感的にも妥当な傾向が獲得されている。

また構築した図3のモデルを用いて、レストラン推薦システムを試作した(図4)。ユーザ変数、状況変数より、好みのコンテンツ属性を確率分布として予測する。これとコンテンツの属性の値を対数尤度により評価し、次のスコアを求める。

$$A_i = \sum_{j=1}^n \log p(c_j = C_{ij}) \quad (3)$$

このスコアの値が高いコンテンツ順に、推薦することで状況とユーザに適応したカーナビが実現できる。この試作システムと従来のカーナビと比較したところ推薦結果のユーザの好みと状況にあったレストランが上位に提示される点で

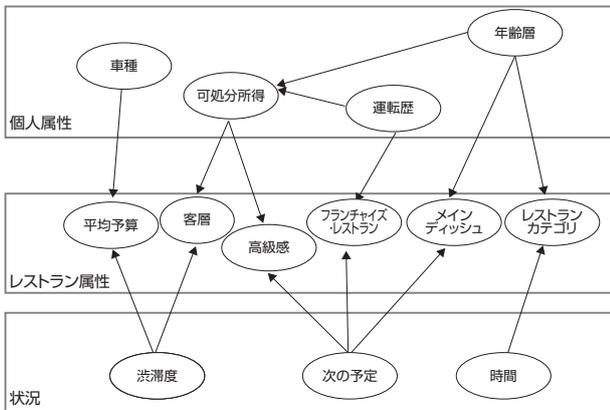


図3 レストラン嗜好ベイジアンネットワークモデル

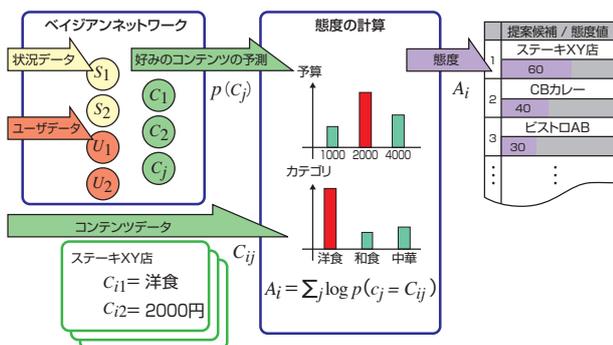


図4 レストラン推薦システムの概要^[6]

有効性が確認された^[6]。

7.2 携帯電話によるユーザ・状況適応型情報推奨

次世代の携帯電話サービスにおいても、多様なユーザや状況に適応する情報推奨技術は重要である。ここでは携帯電話サービスにおける映画推奨サービスにベイジアンネットワークを適用した事例^{[19][20]}を紹介する。まず、約1600名の被験者に対して映画コンテンツを提示するアンケート調査によりユーザ属性、コンテンツ属性、コンテンツ評価履歴を取得した。年齢・性別・職業などのデモグラフィック属性の他にライフスタイルなどに関する質問項目、さらに映画視聴に関する態度属性として鑑賞頻度、映画選択時の重視項目、映画を見る主要目的(感動したい等7項目)、コンテンツに対する評価{良い・悪い}、その時の気分(感動した等7項目)などを収集した。さらに約1000人について別途、各映画コンテンツについて、どんな気持ちや状況で、どこで(映画館、DVDで家)、誰と何人で、どんな時に、鑑賞するか、を自由記述文により収集した。このデータを筆者が開発したベイジアンネットワーク構築ソフトウェア BayoNet^{[17][18]}に入力し、自動的にベイジアンネットワークモデルを構築した。

こうして構築したベイジアンネットワークにより状況とユーザの嗜好性に応じて映画を推薦する携帯情報システムのプロトタイプを開発した。ユーザが携帯電話からサービスへの要求を状況に関する情報とともに送ると、システムはデータベースから登録済みのユーザ属性情報と状況情報を使って確率推論を実行する。その結果選択される確率が高いと判断されたコンテンツを上位から推薦する(図5)。この映画推薦システムはインターネットサービスにも発展し、auoneラボ(<http://labs.auone.jp>)において2007年から一般に公開されるべ約7000件の推薦を実行した。その推薦履歴からさらにモデルの再学習を行うことで推薦精度を向上させる実験も行っている。またこのように構築された映画選択の計算モデルを用いて、今度は映画公開が終わったコンテンツのDVD販売戦略の最適化を映画配給会社と共同

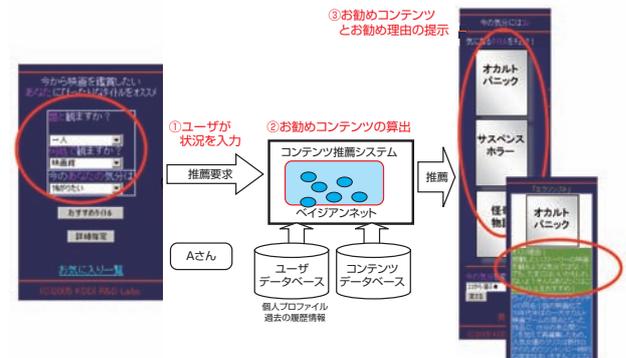


図5 ユーザと状況に応じて映画を推奨する携帯情報サービスシステム^{[19][20]}

で推進している^[21]。

こうした情報サービスが普及し、多数のユーザがシステムを利用することによって、選択したコンテンツの履歴がさらに大量の統計データとして集積する。そのデータによりベイジアンネットモデルの改善が進み、モデルの適合度や推論精度も向上するといった好循環と別のサービスへの水平展開も実現できる。これは実サービスを通じて市場から得られるデータが計算モデルにより再利用可能な知識となり、これがさらに次のサービスに反映される知識循環、すなわち先に述べた、Research as a Serviceの好例と言える（図6）。このような実サービスを通じた研究活動はサービス工学研究センターにおいても実践され、大規模データからの計算モデル構築によって、最適設計ループをフィールドの中に実装し、サービス産業の生産性を向上させるための事業として推進されている^[22]。

8 おわりに

本研究においてソフトウェアの開発までは第一種基礎研究的であったが、ソフトウェアの開発が一段落した後は明らかにアウトカムを志向した第二種基礎研究を意識したものになった。その過程で直感的に進めてきたアプリケーションの選択基準を今あらためて考えてみるといくつかの条件があったように思う。

1. 既存手法では解決できていない問題であること
2. 利用者が必要としている顕在化している問題であること
3. その問題解決により利益を得、それに見合うコストとリスクを負担するステークホルダーが存在すること

このような条件の中で、ベイジアンネットを人間行動のモデル化に用いて、顧客やユーザの行動を予測すること、それによりサービスを最適化することで価値の向上と効率性の向上を果たすことが適切なアウトカムであると考えられ

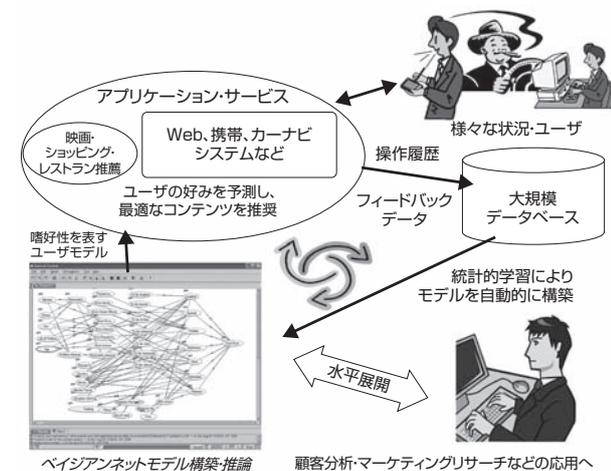


図6 ベイジアンネットによる知識循環サービス

た。このアウトカムを実現できるクライアント企業は多様な顧客との接点（チャネル）を持つ業種の中に存在した。先述したアウトカムを選択したことで、適用フィールドはインターネット、携帯電話、カーナビ、コールセンターなどの顧客からの大量データが集積できるチャネルであることになる。しかし、これらの選択肢の中で、現時点の技術を移転することで十分対応できるものと、アウトカム実現のためにさらなる技術開発が必要なものと2種類があった。そこで前者にはベンチャーが対応し、後者は産総研と企業との共同研究を進めるという選択をした。

工学的な実現と社会的な実現は異なる。工学的にはすでに確立した技術であっても、社会的な価値を生み出すためにはさらに多くのステークホルダーの関与を必要とする。必ずしも工学的バックグラウンドを持つわけではないこうしたステークホルダーにアウトカムの価値を伝え、コストとリスクを担ってもらうためには、アウトカムの効果を信頼性の高い形で示すことが必要であった。そのために事業部レベルでの共同研究や産総研技術移転ベンチャーを通じた社会実装が必要となり、効果は実フィールドで実証される必然性を持つ。つまりアウトカムの評価と社会実装は同時に行われることになったのである。

社会的実装が可能な条件を明らかにするために、ベンチャータスクフォースの中で市場調査を行った。そこでは第一種基礎研究では考える必要のなかったコストベネフィット分析が重要であった。社会実装を円滑に進めるためにはベネフィットの向上とともにコストとリスクの低減が求められる。この段階でアウトカム自体の修正、あるいは新たなアウトカムが必要となることから、基礎研究の駆動力となる可能性がある。これこそが本格研究を推進する上で新たに生まれる第一種基礎研究へのフィードバックであり、デジタルヒューマン研究センターのポリシーステートメントとして

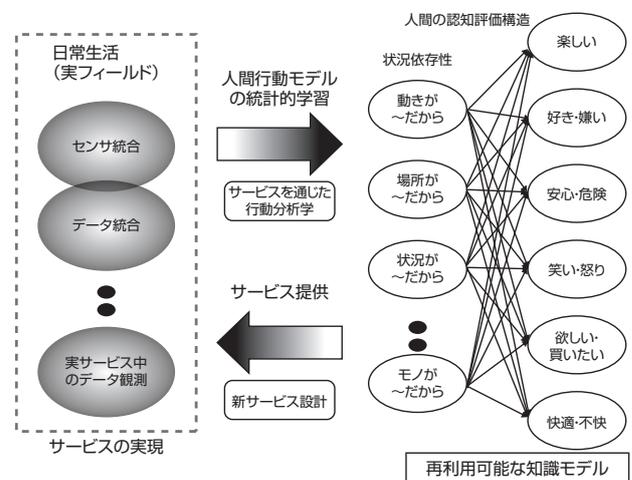


図7 再利用可能な人間の認知・評価構造モデル

掲げられている「アプリケーションに駆動された基礎研究」ということになる。そして実サービスを通じた活動の結果、実ユーザをとりまく状況や文脈までも含んだ大規模データを獲得することが可能になった。またこのデータから構築したベイジアンネットは実在する消費者、生活者の認知・評価構造や行動を予測し、データの記述モデルにとどまらない因果的モデルであることにより、これは他のサービスにも水平展開可能な再利用性の高い知識モデル（図7）として集積し活用できるものになる^[23]。

社会実装のために要請される課題に対してどのように基礎研究の立場で迅速な回答が生み出せるかが、これからの社会から要請される問題解決型の基礎研究を確立する際の重要課題であると考えられる。社会技術としてのスピードが求められるからこそ、あらかじめ多くの芽を養っておく必要がある。その選択は基礎研究に通暁した者にしか行えないが故に、基礎的な研究を行うためには将来を見据えた社会技術を志向する見識が強く求められるであろう。

参考文献

- [1] S. Russell and P. Norvig(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 (2003).
- [2] J. Pearl: *Probabilistic inference and expert systems*, Morgan Kaufmann, CA, (1988).
- [3] P. Baldi, P. Frasconi and P. Smyth: *Modeling the internet and the web – probabilistic methods and algorithms*, [確率モデルによるWebデータ解析法, 森北出版 (2007)].
- [4] D. Marr: *Vision: A computational investigation into the human representation and processing of visual information*, W.H. Freeman and Company (1982).
- [5] G. Cooper and E. Herskovits: A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9(4), 309-347 (1992).
- [6] 本村陽一, 岩崎弘利: *ベイジアンネット技術*, 東京電機大学出版局 (2006).
- [7] Y. Motomura and T. Kanade: Probabilistic human modeling based on personal construct theory. *Journal of Robotics and Mechatronics*, 17 (6), 689-696 (2005).
- [8] 本村陽一, 西田佳史: 日常環境における支援技術のための行動理解. *人工知能学会誌*, 20 (5), 587-594 (2005).
- [9] 美濃導彦: ユビキタスホームにおける生活支援. *人工知能学会誌*, 20 (5), 579-586 (2005).
- [10] B. F. Skinner: *Behavior of Organisms*, Appleton-Century-Crofts (1938).
- [11] 白石康星, 保川悠一郎, 西田佳史, 本村陽一, 溝口博: 日常生活行動情報収集管理システム. *人工知能学会全国大会*, 3G3-03 (2008).
- [12] Y. Nishida, Y. Motomura, G. Kawakami, N. Matsumoto and H. Mizoguchi: Spatio-tempora semantic map for acquiring and retargeting knowledge on everyday life behavior. *Lecture Notes in Artificial Intelligence, JSAI 2007 Conference and Workshops, Revised Selected papers*, 63-75, Springer-Verlag (2008).
- [13] 川上悟郎, 西田佳史, 本村陽一, 溝口博: ロケーションEMGセンサを用いた行動の時空間展開記述に基づく日常生活行動モデリング手法. *知能情報ファジィ学会誌*, 20 (2), 190-200 (2008).
- [14] 石川詔三, 河田論志, 本村陽一, 西田佳史, 原一之: 日常生活行動における確率的因果構造モデルの構築と行動推定. *人工知能学会全国大会*, 3G3-04 (2008).
- [15] 本村陽一, 西田佳史: ベイズ推定における事前分布のグラフ構造モデリングと実生活行動理解. *情報処理学会論文誌コンピュータビジョンとイメージメディアCVIM*, 18, 43-56 (2007).
- [16] 科学技術振興事業団, 科学技術未来戦略ワークショップ (電子情報通信系俯瞰WS) 報告書 (2007).
- [17] Y. Motomura: BAYONET: Bayesian network on neural network. *Foundation of Real-World Intelligence*, 28-37, CSLI california, (2001).
- [18] 本村陽一: ベイジアンネットソフトウェアBayoNet. *計測と制御*, 42 (8), 693-694 (2003).
- [19] C. Ono, M. Kurokawa, Y. Motomura and H. Asoh: A context-aware movie preference model using a Bayesian network for recommendation and promotion. *Proc. of User Modeling 2007, LNCS*, 4511, 257-266, Springer, (2007).
- [20] 小野智弘, 本村陽一, 麻生英樹: 移動端末におけるユーザの状況を考慮した嗜好抽出技術. *情報処理*, 48 (9), 989-994 (2007).
- [21] 落合, 下角, 小野, 麻生, 本村: ベイジアンネットワークを用いた映画コンテンツのマーケティング支援. *人工知能学会全国大会* (2009). (投稿予定)
- [22] 本村陽一他: サービスイノベーションのための大規模データの観測・モデリング・サービス設計・適用のスパイラル. *人工知能学会誌*, 23 (6), 736-742 (2008).
- [23] 本村陽一, 西田佳史: 人間行動理解研究はなぜ難しいのか? ~研究を加速するための知識共有システム~, *人工知能学会全国大会*, (2007).

執筆者略歴

本村 陽一

1993年電気通信大学大学院博士前期課程修了。同年工業技術院電子技術総合研究所入所、2000年情報科学部主任研究官、2001年産業技術総合研究所情報処理研究部門主任研究員、2003年～同研究所デジタルヒューマン研究センター主任研究員、2008年～同研究所サービス工学研究センター大規模データモデリング研究チーム長兼任。博士(工学)。モデライズ(株)取締役兼CTO。人工知能学会全国大会優秀賞、研究奨励賞、ドコモモバイルサイエンス賞など受賞。電子情報通信学会、日本神経回路学会、日本認知科学会、日本行動計量学会、マーケティングサイエンス学会、IEEE各会員。

査読者との議論

議論1 構成学としての独創性を明瞭に記述することについて

質問・コメント(中島 秀之)

論文(第1稿)の第2章前半に日常生活行動をモデル化することの困難性と統計学習によるアプローチの重要性が書かれていますが、専門外の人には分かりにくいので、もう少し具体的な話題に展開していただく方が分野外の人にもわかりやすいと思います。また第2章の後半に、「サービスとしての調査・研究(Research as a service)」が書かれていますが、これがこの論文の本質でしょう。この部分を膨らませてください。

質問・コメント(持丸 正明)

本研究はブレイクスルー型の本格研究と見ることはできないでしょうか(図a)。ベイジアンネットワーク技術が中核となる重要技術要素で、それにセンシング技術やインタビュー技術などが周辺技術として統合され、実社会の問題解決につながったという“構成学”であると考えられます。ここで“構成学”として特に興味深い点は、単に周辺技術を統合するだけではこの研究は完遂しないという点です。

そのために、「実社会で有効なサービスを実現しながら、それを通じて調査・研究を遂行する」という新しい“構成学”の枠組みを提案している点に、“構成学”としての独創性があると思います。社会循環型の本格研究と言えるでしょうか（図 b）。過去に Synthesiology 誌に掲載された論文にも、この社会循環型の本格研究の類型に当てはまりそうなものもありますが、やはり、本論文がもっとも明瞭にそれを体現していると思います。“構成学”の学術誌でもあるので、やはり、この独創性の部分をより明瞭に記載してください。

回答（本村 陽一）

ご指摘の点を踏まえ、アブストラクトにおいて各要素技術とその構成としての本研究のポイントを明記し、題目、位置づけを Research as a Service を主体として修正しました。

議論2 論文の構成（章立て）について

質問・コメント（持丸 正明）

人工知能学会誌の論文ではなく、構成学の論文であることを考えると、導入部分は統合された技術がもたらす「夢」であるべきと思います。「人間がなんの目的を持って行動するかを理解するシステムによって人間生活を支援するサービスを実現する」という夢とその具体的な事例イメージを、読者に最初に提示するのがよいと思います。その夢を実現するためのブレイクスルーポイントが「日常生活を計算論的に、記述・理解・実現する」ということになるでしょう。それをブレイクスルーするための難しさとして、(1) 人間の行動という曖昧で不確実な要素を含んでいる点があります。それを乗り越えるための方策として非決定論的枠組みが有効であり、具体的技術としてベイジアンネットワークを適用したと言うことになるでしょう。また、ベイジアンネットワーク技術によって派生する難しさとして、(2) 大量データの観測というものがあります。これをユビキタスセンシングや実社会でのサービス（RaaS）によって解決していくという筋道だと理解しています。そのような章立てに直した方が読みやすいと思います。

質問・コメント（中島 秀之）

「1. はじめに」の最後の部分にそれ以降のあらすじを記述していたのが良いかと思います。第2章（RaaS）が唐突に書かれていて、その後うまくつながっていません。

質問・コメント（持丸 正明）

副査読者の中島氏からも指摘があるとおりの、第2章（RaaS）の配置がスムーズな論旨展開の妨げになっているようです。主査として3つの解決策を提案します。第一は、副査読者の中島氏が提案しておられるように、章立ては変えず、第1章（はじめに）の末尾に本論文の論旨展開を記載するという事です。第二は、第2章（RaaS）と第3章（非決定論的アプローチ）の順番を入れ替えるというものです。第三は、第2章（RaaS）を思い切って、第8章（おわりに）の前に持ってくるという事です。第一の「論旨展開」を第1章の末尾に書くという事は、章立ての順番を入れ替えることとは独立の方策ですので、いずれにせよ、実践いただくのが読者のためによいかと思います。

主査としては第三の解決策を推奨します。本論文では「人間行動モデルに非決定論的な方法論を選択」「非決定論的な方法論としてのベ

イジアンネットワーク」に加えて「RaaS」が書かれています。具体的な事例を提示する前に、これらの理論（あるいは考え方）を一斉に読者に示しても、受け止めきれないのではないかと懸念します。そこで、前半では「非決定論的なアプローチ」を中心に論旨を展開し、具体的な事例を通じて、それを実現するのに意味のある大量データが必要であることに触れ、それを取得して研究を推進する方法として「RaaS」を提案して、さらにそれを展開するためにステークホルダーとの連携形成が不可欠である、という流れにしたいかどうか、ということです。

回答（本村 陽一）

ご指摘どうもありがとうございます。いただきましたコメントに基づきまして、章立てを第三の解決策に従って修正いたしました。

議論3 具体事例に関する記述の追加について

質問・コメント（中島 秀之）

「こうした情報サービスが普及し、多数のユーザがシステムを利用することによって、選択したコンテンツの履歴がさらに大量の統計データとして集積する。そのデータによりベイジアンネットワークモデルの改善が進み、モデルの適合度や推論精度も向上するといった好循環が実現できる。」という部分がこの論文のもっとも肝要なポイントであると思います。もう少し具体的にどのようなスパイラルになったのかを記述して下さい。1000名以上を対象に実証実験を行ったことは評価しますが、ちょっときつい言い方かもしれませんが、これだけでは実用に供したとは言えません。社会科学ではこれくらいの数のアンケートは普通ですし、システムの試作と実証実験までは従来より情報系でも行われてきたことです。そこで止まらず、実際のサービス提供をするというフェーズが大事なのだと思います。

回答（本村陽一）

ご指摘の点を補強して、事例に関する情報を追記いたしました。

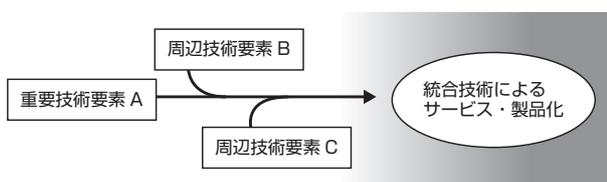
議論4 再利用可能なモデルの説明について

質問・コメント（持丸 正明）

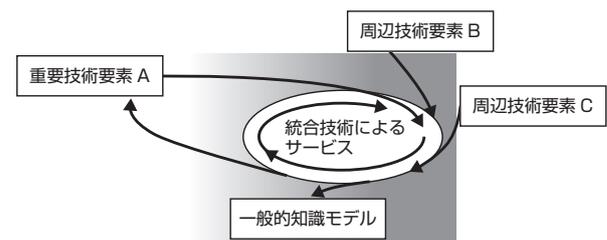
「再利用可能なモデル」の具体的なイメージを示している図（最終稿の図7）は、査読者がコメント1で提示した当該論文の“構成学”の枠組みの図（図b）における「一般的知識モデル」に相当するものかと思えます。サービスを実施しながら調査・研究を進めることで、他のアプリケーション（サービス）に水平展開できるモデルができるというのが、この論文の提案・実証する“構成学”としての面白味であります。一方で、実社会でのサービス循環を通じて研究を進めるということに加えて、「再利用可能なモデル」が生成されるというのは、読者にとってはやや情報量が多く、即座に理解しにくいのではないかと懸念します。そこで、論文前半ではこの点にさらっと触れるのみにして、事例を示したあと、第8章（おわりに）のところで図を提示して、改めて「再利用可能なモデル」が生成されることを詳細に述べる方が効果的ではないかと考えます。

回答（本村 陽一）

ご指摘の通りに図の移動と第2章、第8章を修正いたしました。



図a ブレイクスルー型



図b 社会循環型