

ベイジアンネットによる身体知のモデル化

Modeling Physical Tacit Knowledge by Bayesian Networks

古川康一[†] 植野研^{*} 五十嵐創[†] 森田想平[†]
尾崎知伸^{††} 玉川直世^{††} 奥山渡志也^{†††} 小林郁夫[†]

[†] 慶應義塾大学大学院政策・メディア研究科 ^{*} 東芝研究開発センター

^{††} 慶應義塾大学 SFC 研究所

^{†††} 慶應義塾大学総合政策学部

furukawa@sfc.keio.ac.jp

<http://bruch.sfc.keio.ac.jp/>

Abstract: 身体知は、訓練によって獲得される、筋骨格系の整合的な一連の動作と考えられる。身体知獲得の目的は、演奏、踊り、各種のスポーツなどのスキルを向上させることである。身体知の問題は、それが暗黙的であり、職業演奏家や、プロスポーツプレイヤーが自身で何を行っているのかを把握出来ない点である。本研究の目的は、身体知をモデル化し、その暗黙知を言語化することである。そのモデル化に、ベイジアンネットを用いることを検討した。ここでは、その基本構想を明らかにする。

1 はじめに

ヒトが知的能力を発揮する場面は、言葉の理解に始まる論理的な側面にとどまらず、楽器の演奏、ダンス、スポーツなど、多岐にわたる身体能力の発現の場にも広く観察される。それは、動物の能力と比較することにより、明らかである。運動の身体能力だけに限れば、ヒトが他の動物に劣る点も数多く見受けられる。最も典型的な例は、鳥との比較である。鳥は飛べるが、人は飛べない。しかしながら、より複雑な動作になると、その優劣は明らかである。それらは、楽器の演奏、ダンス、各種のスポーツなどであり、それらが本論文で取り上げる身体知が関与する動作の対象である。

暗黙知を獲得するためには、正しい練習法が必要不可欠である。弦楽器を例にとると、各種の練習法が知られている。鈴木メソッド、桐朋スクール、ロシアスクールなどは、その例である。このような練習法は、経験から生まれた、探索空間の発見的な枝刈り法となっている。バイオリンの奏法を例にとると「最適」な奏法は、時代とともに、あるいは、流派によって異なる。弓を保持する腕は、ロシア派では、上腕を肩より上に構えることが推奨された。近代奏法では、これは否定されている。チェロの奏法の例では、これまでチェロをやや斜めにして、楽器の胴が身体を中心に来るように構えるのが習慣であったが、Victor Sazer は、チェロをもっと身体の左に寄せて、方向は正面から見て真っ直ぐになるようにした方が良く、と主張している [Sazer 95]。また、彼はチェロを弾く際の最適な椅子の形について、後ろを数センチ高くするという大変ユニークな提案をしている。このように、最適解の発見への努力がいまだに続いているとすることができるであろう。

暗黙知のモデル化、あるいは言語化は、このような、経験的に得られた練習法に対して科学的裏付けを提供する有力な方法である。練習法によって獲得される技能が、暗黙知の言語化によって明らかにされれば、各練習ステップの位置づけが明確になり、かつ、生徒にとっても納得のいくものとなるであろう。

身体知のモデル化、ないしは言語化の試みは、いくつか見られる。その第1は、仮説検証型のアプローチで、それは、うまい演奏を可能にする要因を列挙し、各要因について仮説を立て、実験によりその仮説を検証する方法である。たとえば、渋谷らは、バイオリンの奏法に関して、熟達者と初心者の弓の使い方を分析し、熟達者は、弓の位置によって圧力が大きく変化しないこと、および弓を返す時に熟達者の方が指の動きが大きいことを確かめた [渋谷 94]。

第2は、熟練者およびアマチュアのパフォーマンスのデータから、データマイニングの技術を用いて、熟練者とアマチュアの違いを明らかにする手法である。我々は、これまで、チェロの演奏について、決定木や帰納論理プログラミングを用いて、これらの差を抽出する研究を行ってきた [Ueno 00, Furukawa 99, Ueno 98]。

本研究は、これまでの研究のアプローチでの困難性を解消することを目的としている。これまで、我々は、上の第2のアプローチにより、チェロの運弓動作の分析を進めてきた。そこで、上級者とアマチュアの違いを抽出することは出来たが、その明確なルール化にはいまだ成功していない。その主な原因は、筋肉群の使い方の違いが影響していると思われるが、それが明確に観測出来ないことと、それらが確率的に変動していると思われる点である。

一方、楽器の演奏は、発話、あるいは朗読とよく似たプロセスであることが分かる。楽譜を見ながらの演奏を考えると、楽譜の指示通りに手や指を動かし、楽譜が要求する音楽を生成する。それに対するのは朗読で、書いてある通りに読む。暗譜による演奏は、演劇での台詞の発話に相当し、ジャズなどでの即興演奏は、より自由な環境での発話に相当すると考えられる。これらのアナロジーは、発話の解析に隠れマルコフモデル (HMM)[中川 02] やベイジアンネット [Russell 95] を用いることから、楽器演奏の解析に対しても同様の手法を適用することの妥当性を示唆している。より具体的には、演奏の場合、あるプランに従って筋肉群を動かし、曲を演奏するが、一方、発話の場合には、発話の意図に従って口の周りの筋肉群を動かし、発話する。

一方、演奏と言語の発話との相違点も注意すべきである。発話の場合にベイジアンネットのノードに置かれる音素に対応するものが、運動スキル、あるいは演奏の場合、明確ではない。このことは、HMM やベイジアンネットでのモデル化の際に、問題をどのように表現すべきかに関わってくる。また、モデル化によって何を行うのかも、重要な問題である。発話の場合、その目的は予測による発話理解であるが、演奏の場合、予測が問題にされるわけではない。

本論文では、これらの問題を以下のように順次取り上げ、議論をしていきたい。2節では、モデル化の目的について述べる。3節では、ベイジアンネットによるモデル化について論じる。4節では、観測データの利用と、条件確率の計算について論じる。5節では、まとめと今後の課題を与える。

2 モデル化の目的

暗黙知のモデル化の目的は、各動作の認識 [Ioffe 01]、あるいは予測を行うためではない。その点が、HMM による発話認識と大きく異なる。その目的は、ある水準に達した熟練者の演奏モデルとアマチュアのそれを別々に構築し、その相違を明らかにすることである。あるいは、熟練者の通常時の演奏モデルと、時たま陥るスランプ時の演奏モデルとを比較して、スランプの原因を特定するのにも使いたい。そのためには、ベイジアンネットの構築問題に取り組みなければならない。また、モデル間の相違を主張するためには、モデル自身の信頼性、およびモデル間の差の有意差検定などを論じるための方法論を必要とする。

データマイニングによる暗黙知のモデル化の一手法に「行動のクローン化」(behavioral cloning)[Sammut 92] が知られているが、そのターゲットは熟練者の行動を真似るクローンプログラムの生成である。それは、ルールの形で得られる。もし、そのルールが適当なロボットの制御プログラムとして組み込まれれば、まさに行動を真似るロボットが実現できることになる。実際には、熟練者の行動パターンを示す定性モデルの構築を行う。

本研究の狙いの一つは、このような定性モデルと、ここで展開しているベイジアンネットによる確率モデルを統合して、定性確率モデルを構築することである。定性モデルは、物理システムの定性的な変化を記述するモデルであり、物理システム内の構成要素間の物理的因果関係に基づいて、その変化の可能なパターンを記述する。一方、ベイジアンネットが表現する結合確率分布は、必ずしも物理的因果関係に基づく関連だけを扱うものではないが、ヒトの四肢のような物理系の各構成要素間の事象は、物理的因果関係としても解釈できる。この両面を同時にモデル化する手法を目指したい。

3 ベイジアンネットによるモデル化

ベイジアンネットによる身体知のモデル化の表現を決めるためには、ベイジアンネットの各ノードにどのような事象を置けばよいのかを決め、さらに、ノード間の相互関係を決めなければならない。また、各ノードの条件確率表の確率値を決定する必要がある。ここでは、その中で、前二者について論じる。表現を決める上で問題となるのは、ノードに置かれる事象の複雑さである。もし、それらを命題記号に限ると、その表現は単純になる。しかし、その

場合、たとえば二つの事象の相関が非常に強いとすると、それら二つの事象の連言をノードに置いた方が良いかもしれない。この選択により、以下の二つの方式が考えられる。

方法1．ノードに命題論理のアトムを置く．そして、筋骨格系に沿った構造を作る．

方法2．ノードに命題論理のリテラルの連言を置く．各リテラルは、背中から指に至る筋骨格系の力学的状態を表す．

前者は関節およびそれらにつながっている筋肉群についての特定の状態をノードとし、それらの筋骨格系に沿った関連をリンクとするので筋骨格系ベイジアンネットとよび、後者はノードに置かれるリテラルの連言がある時点におけるモデルを表すので、モデル遷移ベイジアンネットと呼ぶことにする．

以下に、この二つの方式のそれぞれについて、検討する．

3.1 筋骨格系ベイジアンネットによる表現

ノードに置かれるアトムは、腰から上の各関節の角速度、角加速度の状態、および筋肉群の活動状態である。たとえば、肩関節を考えると(1)腕を横に上げる(2)腕を前に上げる(3)腕を腕方向に沿って内側、あるいは外側に回転させる、の3種の独立した回転がある。また、肘の場合は、角度の変化は一通りしか起こらない。もちろん、それらの状態は時間とともに変化するので、時間変化の表現が必要である。時間的な変化は、HMMでの時間軸に沿ったネットワークの構成の方法と、ある時刻でのベイジアンネットワークから次の時間のそれへと繋げていく、動的ベイジアンネットワークの手法が知られている。音声認識でのHMMは、音素の系列が時間軸に沿って並ぶ構造を持っているが、ここでは、腰から指に至る一連の動作の伝播が時間軸上に並ぶことになる。また、ある関節の動きは、必ずしも直前の筋肉群のみによって決まらず、それ以前の筋肉のモードに影響を受けるので、単純マルコフ性だけでは問題を表現出来ない。たとえば、手首に対しては、前腕の筋肉群だけでなく上腕のそれも影響を及ぼしていることが知られている。そのため、それを越えたベイジアンネットの表現能力を必要とする。また、弓の動きで、下げ弓、上げ弓を繰り返す運動をこのベイジアンネットで表現しようとする、再び同じ構造のネットに戻り、それが繰り返されることになる。このような構造は、動的ベイジアンネットに見られる。ここでは、この両者の表現を組み合わせる必要がある。ところが、もう一つの問題は、動的ベイジアンネットのこれまでの使われ方が、次々に送られてくるセンサー情報をモデル化している点である。すなわち、それらのシステムが相手にしているのが、状況に応じて反応するようリアクティブな動きであり、複雑なプランに沿った動きは想定していない点である。そのため、ここでの目的に沿った、より複雑な運動制御を表現するための動的ベイジアンネットを構築する必要がある。

もう一つの問題は、関節角と筋肉の活動の二つに、観測上の差異がある点である。各関節角の速度、加速度は、モーションキャプチャリングシステムによって得られるデータから、比較的容易に計算出来るが、各筋肉の活動レベルに関しては、表面筋電図がセンサー周辺の筋肉群の活動を加え合せたものしか計測出来ない。そのノードでの条件確率表を直接推定することが出来ない。そのため、ここでの筋骨格系ベイジアンネットでは、筋肉に対応するノードを隠れ層とみなす。そのようにして得られた、ベイジアンネットを図1に示す。ただし、この図では、動的ベイジアンネットは含まれていない。そのイメージは、この図全体が多少の変更を加えつつ繰り返される、という構造となる。

3.2 モデル遷移ベイジアンネットによる表現

モデル遷移ベイジアンネットは、リテラルの連言を表現言語とする方法である。各ノードは、ある時点で成り立っている事実を表すと考えられるので、モデルと考えてよい。すなわち、ここでのベイジアンネットは、様相論理での可能世界モデルに対応する。この表現の利点は、同時に成り立っている事柄を一つのノードで表現できる点である。たとえば、拮抗筋の利用とは、屈筋と伸筋の同時使用のことであるが、本モデルでは、それは屈筋使用および伸筋使用が同時に成り立っているノードとして、自然に表現できる。他の例としては、上腕と前腕から手首までを

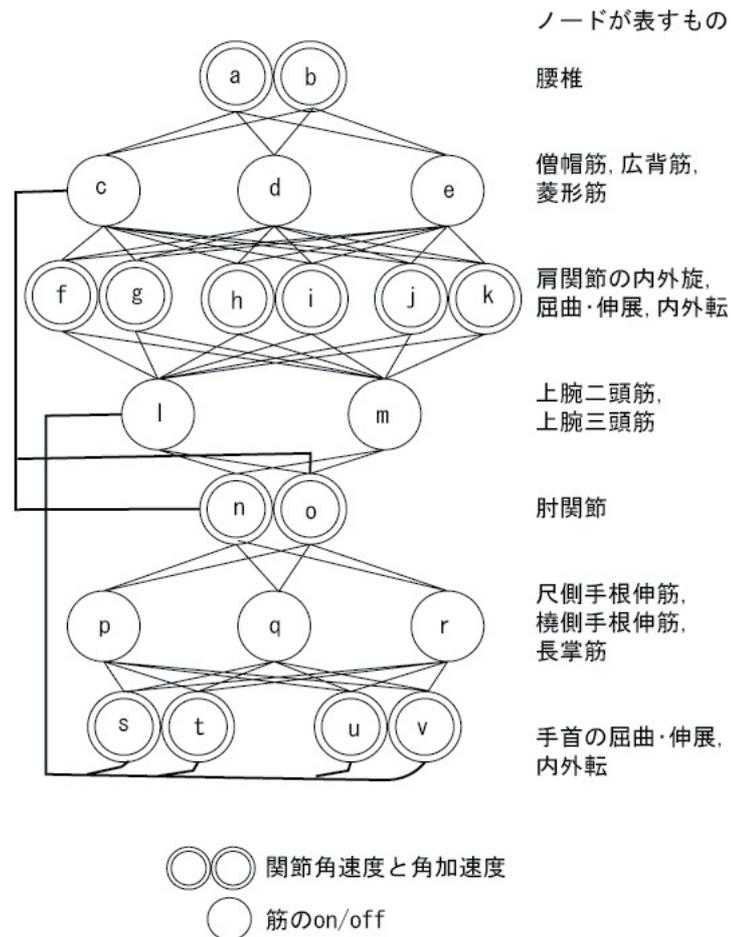


図 1: 筋骨格系ベイジアンネット

一体として動かすモードと鞭のように動きを伝播させるモードの区別がリテラルの連言で表現できるので、便利である。本方法での時系列の扱いは、あるノードから他のノードへの遷移として、表現できる。

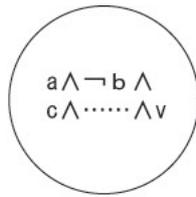
本方式の表現例を図2に示す。図2(a)では、ノードの表現例を示す。また、図2(b)では、「反射」、および「拮抗の解除」の2例でのノード間の遷移を示す。

本方法の問題点は、もともとの結合確率モデルの設定である。各ノードのリテラルに時刻によるインデックス付けをすることにより、時空間での結合事象を考えることが出来る。また、各ノードがモデルを表現するためには、フレーム問題が発生するので、その回避も問題となる。そのために、状況計算などの手法を導入する必要性が生じてくる。表現力が豊かな分、そのための推論メカニズムが複雑になると考えられる。

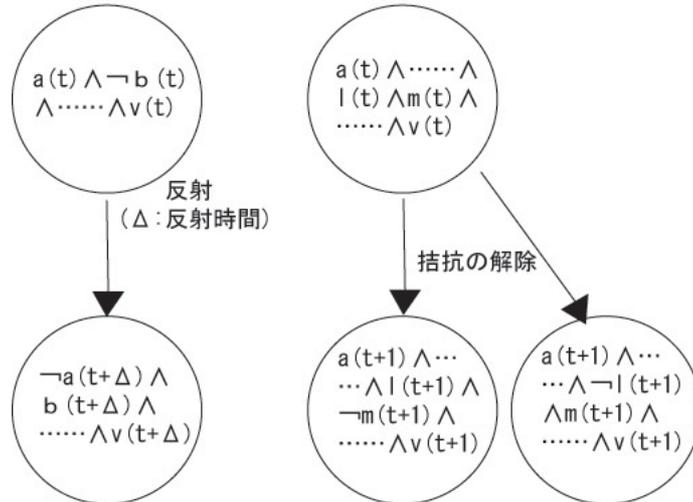
4 観測データに基づく条件確率の割り当て

観測データとしては、モーションキャプチャリングデータ、筋電図データ、楽器に取り付けられた圧力センサーからのデータなどがある。この中で、モーションキャプチャリングデータから、身体の各関節部分の位置、角速度、角加速度の情報が得られる。実際には、時系列データに対して、適当な前処理を施して、セグメント化をしなければならない。セグメント化については、現在、新たなアルゴリズムを開発中であり、別の機会に報告する予定である。これらの観測データから、各状態間の遷移確率を推定できる。これらの遷移確率は、ベイジアンネットの関節ノードの条件確率を推定するのに利用できる。すなわち、関節ノードの各条件ごとに、その遷移の発生する頻度を集計すればよい。

厄介なのは、筋電図の扱いである。とくに、一般的に利用できるのは表面筋電図であり、それは、上で述べたよ



(a) ノードの表現例



(b) ノード間の遷移の例

図 2: モデル遷移ベイジアンネット

うに、各筋肉の活動のみを計測できるわけではなく、周りの筋肉群の活動も加わった、誤差の多いデータしか得られない。この問題を解決する一つの方法は、センサー毎に、関連する筋肉群を含めてモデル化し、各筋肉の成分を、たとえば別のベイジアンネットにより推定する方法である。もし、各筋肉を独立に動かしたときの動きが分かれば、このような手法も有効であろう。もう一つの問題は、筋電図が力の絶対値を測定出来ない点である。また、それに伴い、各測定値間の力の比も調べることが出来ない。できるのは、測定値ごとの変化の観測である。本研究では、そのため、筋肉については、オン、オフの2値、あるいは、中間的な力を含めた数段階の多値による離散化を考え、定性推論の対象とすることを考える。

仮説としては、各筋肉の使用モードを想定しているので、最終的に計算したいのは、 $\Pr(\text{拮抗筋オン} \mid \text{手首の切り替えしオン})$ 、あるいは、 $\Pr(\text{拮抗筋オン})$ などの確率である。さらに、アマチュアとプロの演奏家に対するベイジアンネットの違いから、それらの確率値を求めて、具体的な違いを計算できることになれば、有用であろう。

5 おわりに

本論文では、ベイジアンネットによる身体知のモデル化の基本構想を明らかにした。そして、身体知のモデル化にとって、ベイジアンネットが十分な表現力を持っていることを確認した。

次のステップは、観測データから、実際に各ノードにおける条件確率表を求め、ベイジアンネットによって、さまざまな事象の確率値を予測し、それによって、種々の仮定を検証することである。このためには、すでに手持ちのデータの他にどのようなデータが必要かを明らかにしなければならない。また、EM アルゴリズムや、MCMC アルゴリズムなどによって、どこまで未知の確率値を予測できるのかを調べなければならない。

具体的には、本モデル化を通して、弓の返しの複雑な動きのモデル化を実現したい。そのエッセンスは、下げ弓

と上げ弓で、最適な筋骨格系のポジションが異なる、という事実である。いずれも手首によって先導されることには変わらないが、下げ弓の場合、手首の小指側で手を引っ張るように動かし、上げ弓の場合には、手首の親指側で手を引っ張るようにする。このため、前腕と手のなす角度が、下げ弓と上げ弓で異なる。この、前腕と手の最適な配置を、弓の返し時に瞬間的に達成しなければならない。より厳密に言えば、弓の方向変換に先立って、この手の形の変化を達成しなければならない。このためにチェリストが行っていることは、手首の返しである。手首を瞬間的に返すわけである。物理学的には、この手首の返しは、手首から先の部分の慣性運動によって達成されると考えられる。また、生体における運動制御においては、この瞬時の変化は、脊髄反射によって達成されていると考えられる。このような現象を、より厳密に検証するのが、今後の課題である。

また、より一般的には、基本構想の構築過程で、我々はベイジアンネットと定性推論の関連づけの重要性を認識した。我々の現在の知識では、それらの二つを結びつける提案を見出すことは出来なかった。この新たな研究は、今後の課題である。

さらに、時間軸の扱いを考慮すると、より表現力の大きい、PRISM[Sato 97]などの確率論理プログラミングの導入も検討する価値がある。

生体力学的には、反射運動の利用が暗黙知の獲得に大きく寄与すると考えられるが、本研究により、その点を明らかにする手法を開発するのも、今後の課題である。

参考文献

- [Furukawa 99] Furukawa.K.. "A Framework for Verbalizing Unconscious Knowledge based on Inductive Logic Programming." In K. Furukawa, D. Michie, and S. Muggleton (eds.), Machine Intelligence, 15, Oxford Press. pp.18-24. 1999.
- [Ioffe 01] Ioffe.S. and Forsyth.D.. "Human tracking with mixtures of trees." Proc. of the International Conference on Computer Vision. volume I. pp.690-695. 2001. (<http://citeseer.nj.nec.com/ioffe01human.html>)
- [中川 02] 中川聖一. "音声認識においてHMMとトライグラムを越えるもの." 人工知能学会誌 Vol.17 No.1 pp.35-40. 2002.
- [Sammut 92] Sammut.C., Hurst.S., Kedzier.D., and Michie.D.. "Learning to Fly." In Sleeman,D. and Edwards,P., eds., Proc. of the 9th International Workshop on Machine Learning. pp.385-393. Morgan Kaufmann. 1992.
- [Sato 97] Sato.T. and Kameya.Y., "PRISM: a language for symbolic-statistical modeling." Proc. of IJCAI'97. pp.1330-1335. 1997.
- [Sazer 95] Sazer.V.. "New Directions in Cello Playing." ofnote. 1995. (三本雅俊訳, "新しいチェロ奏法" 音楽之友社, 1998.)
- [渋谷 94] 渋谷恒司, 菅野重樹, 加藤一郎. "バイオリン右腕ボーイング動作におけるスキルの分析." 人間工学 Vol.30 No.6 pp.395-403. 1994.
- [Russell 95] Russell.S. and Norvig.P.. "Artificial Intelligence, A Modern Approach" Prentice Hall. 1995. (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.)
- [Ueno 00] Ueno.K., Furukawa.K., and Bain.K.. "Motor Skill as Dynamic Constraint Satisfaction." Electric Transaction of Artificial Intelligence (ETAI). Linkoping University Electronic Press. 2000. (<http://www.ida.liu.se/ext/epa/ej/etai/2000/011/epapage.html>)
- [Ueno 98] Ueno.K., Furukawa.K., Nagano.M., Asami.T., Yoshida.R., Yoshida.F., and Saito.I.. "Good Posture Improve Cello Performance." Proc. of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE-EMBS98). vol.20 pp.2386-2389. 1998.