Modeling Human Skill in Bayesian Network

Koichi Furukawa† Soh Igarashi† Ken Ueno* Tomonobu Ozaki††
Souhei Morita† Naoyo Tamagawa†† Toshiya Okuyama††† Ikuo Kobayashi†

†Graduate School of Media and Governance, Keio University
*Research and Development Center, Toshiba
††Keio Research Institute at SFC, Keio University
†††Faculty of Policy Management, Keio University
furukawa@sfc.keio.ac.jp
http://bruch.sfc.keio.ac.jp/

Abstract: In this paper, we discuss the problem of modeling human skill in Bayesian network. The purpose of skill modeling is to use the model to improve performances in such activities as playing instruments, dancing, and playing various kinds of sports. The problem of human skill is its tacitness: even professional violinists or cellists do not know how they are playing. The purpose of this research is to model human skill in terms of Bayesian network. This paper gives the basic framework of the research by proposing possible representations and structures of the Bayesian networks for human skill, and by defining the purpose of model usage.

1 Introduction

Human beings show their intelligence not only in logical activities such as language understanding but also in physical activities such as playing musical instruments, dancing and playing sports. This fact is obvious by comparing physical activities shown by animals. If you compare only physical abilities, human beings are worse than animals in many respects. For example, human cannot fly but birds can. However if situations become more complex, human beings show better performance: these include playing musical instruments, dancing, and playing sports, on which we investigate as examples of physical knowledge in this paper.

Continuous and the right practice is the must to get tacit knowledge. For example, there are many schools for teaching the "right" methods in playing the violin: Suzuki method, Toho school, Russian school and so on. The practice methods were invented through experience and they are regarded as the results of heuristic pruning in the vast search space of the ways of playing. In the case of the violin playing, the "best" method in one school differs from that of another school: in Russian school, the upper arm position for holding the bow was thought to be higher than the shoulder. However in recent methodology, this method is denied. In the case of cello playing, the standard method for holding the cello is to put it slightly leaning towards left from the player's side. Recently Victor Sazer [Sazer 95] insists to hold it straightly in order to make the bowing movement at the upper end smoother. He also insists to use a special chair having a slight slope from back to front to give hip muscles more freedom. Thus it can be said that the efforts to find the best solution for the playing are still in progress.

Modeling and/or verbalizing tacit knowledge is one of the most promising approach for providing scientific verification for such playing methods obtained through experience. Through the modeling and/or verbalization of tacit knowledge, skills obtained by continuous practice should be clarified and therefore it would help to order practices from easy to hard, for example. It would help also for the students to understand the meaning of each practice more easily.

There are two approaches to model or verbalize tacit knowledge. The first one is making-hypotheses-and-testing approach where we first propose a hypothesis for better performance and then test the hypothesis through experiments on those features listed in advance. An example of such an approach is by Shibuya et al [Shibuya 94] where they focused their attention on the bowing pressure and the magnitude of fingers' movement and concluded that experienced players tend to keep more or less the same bow pressure and also they use fingers more than amateurs do by analyzing measured data. Note that the feature selection is crucial for the success of this approach. The second approach is data-mining approach where we first measure data on various features played by professional as well as amateur players and then extract rules for distinguishing the former from the latter. We employed this approach in cello performance using such machine learning algorithms as Decision Tree Learner and Inductive Logic Programming [Ueno00, Furukawa99, Ueno98, Igarashi02].

The feature selection is also important in the second approach. However if there are sufficient features to distinguish professionals to amateurs, then feature selection procedure is included in the data-mining process. Another difficulty is uncertainty of input data due to probabilistic variation of muscle usages as well as measurement errors brought by surface electromyogram (EMG). This difficulty suggests to adopt probabilistic representation and reasoning such as HMM or Bayesian network. Furthermore, we notice the similarity between playing musical instrument and speaking and/or reading. The case of playing instruments through reading music notes corresponds to reading. On the other hand, playing by memory corresponds to recitation and improvisation corresponds to free speech. These analogies suggest the application of Hidden Markov Model (HMM) or Bayesian Network to modeling and analyzing performance of musical instrument. However, we should note the difference of playing instruments and speech recognition. In the case of speech recognition, there exist base phones which correspond to states and are put at the hidden layer. On the other hand, there is no analogue to base phones in case of skills. This brings a difficulty in modeling skills in terms of HMM or Bayesian Net. Another difference is the purpose of modeling: for speech modeling, the purpose is to recognize spoken words sequence; on the other hand, for modeling skill, the purpose is identifying behavior to distinguish good performance from bad one.

In this paper, we first state the purpose of modeling in Section 2. In Section 3, we discuss the issue of how to model skills in terms of Bayesian network. In Section 4, we discuss measurements and the usage of the measured data in computing conditional probability. In Section 5, we conclude our current work and give future research issues.

2 Purpose of Skill Modelling

One of the purposes of skill modeling, or in general tacit knowledge modeling, is not to recognize or predict each movement, but to compare skilled artists with amateur players by constructing their respective skill models. We also want to apply the skill model construction to identify the reason of lost form by comparing it with good form. For pursuing these purposes, we need to address the Bayesian network construction issues. Furthermore, we need to address model comparison issues as well as model confidence issues to compare skill models. Model comparison issues contain statistical significant difference testing to assert the model difference.

Another well known approach for skill modeling is so-called behavioral cloning approach based on data mining [Sammut92]. It extracts human expertise as a set of rules. They could be put into a robot as a set of control rules to realize a real robot mimicking expertise behaviors. In most cases, however, only a qualitative model is constructed to explain expertise behaviors.

One of our aims is to combine a qualitative model with a Bayesian network to realize qualitative probabilistic model for skill. A qualitative model describes physical causal relations and can be used to infer possible behaviors by executing qualitative simulation. It is quite natural to utilize probabilistic property to prune unrealistic branch in qualitative simulation. On the other hand, the idea of qualitative modeling can be applied to build a proper Bayesian network by introducing land marks to identify possible events. One possible approach in building Bayesian network is first to infer a proper qualitative model from behavioral data and then convert it into the corresponding Bayesian network.

3 Skill Modelling by Bayesian network

We need to identify a set of events to be put on the nodes and a set of pairwise dependencies among nodes to determine Bayesian network representation of skill. Furthermore we need to cocmpute a conditional probability table for each node. Here we discuss the former two representation issues. One factor for determining the representation is the complexity of an event put on each node. If only propositional atoms are allowed, the node representation becomes simple. However, there may be a case that some of the events have strong correlation. Then it may be better to put their conjunction into a single node. Therefore, there are two possible schemes for designing the skill Bayesian network:

1. Put a propositional atom to each node and construct a Bayesian network having the same structure as musculoskeletal system.

2. Put a conjunction of propositional literals to each node where each literal represents some dynamic regime of the musculoskeletal system.

Since the former Bayesian network represents the musculoskeletal system, we call it as a **musculoskeletal Bayesian network**. On the other hand, since each node in the latter network represents a model which represents the state of the musculoskeletal system at a certain time, we call it as a **model transition Bayesian network**. In the following subsections, we will discuss these two representation schemes in detail.

3.1 Representation by musculoskeletal Bayesian network

Atoms put on each node are angular velocities and angular accelerations of joints from the hip to the wrist, and activity status of each muscles connected to them. Note that there may be more than one pair of an angular velocities and an angular acceleration for each joint. For example, the shoulder has three rotation freedoms: (1) rotation in the sagittal plane, (2) rotation in the frontal plane, and (3) rotation around the upper arm. On the other hand, the elbow has only one freedom of rotation. Another complexity is that they change every mili-second and therefore we need to treat time axis. Two possible ways of treating time are reported: one is to adopt HMM architecture and construct a network along the time axis, and the other is to adopt dynamic Bayesian network architecture and to first build each Bayesian network for each time and then concatenate them. If we adopt HMM architecture, we should note that the muscloskeletal Bayesian network has different structure from that of speech recognition: our Bayesian network has a representation of such an event sequences propagating from the hip to the wrist along the musculoskeletal system, whereas the HMM for speech recognition has a sequence of base phones. Furthermore, there is another difference between them: our musculoskeletal Bayesian network cannot be represented by HMM since a movement of some joint may be influenced by not only muscles directly connected to the joint but also those muscles connected to some joints which are located one or more joints apart. For example, the wrist is affected by not only those muscles connected to forearms but also those at the upper arm. Therefore, HMM is not enough and we need Bayesian network representation. Another complexity occurs when we consider to represent the repeated move of bowing consisting of down-bowing, up-bowing and their repetitions. One possible representation of this repetition is to adopt the dynamic Bayesian model. Another possibility is to adopt a probabilistic logic programming language such as PRISM [Sato95] or SLP[Muggleton96]. Figure 1 shows an image of a musculoskeletal Bayesian network. Note that this figure does not represent repeated bow direction changes. It should be represented by a dynamic Bayesian network repeatedly having similar structures shown here.

Another issue is the observational difference between joint angles and muscle activities. Angular velocities and accelerations of each joint can be computed rather easily from the data obtained through motion capturing system. On the other hand, muscle activity for each muscle cannot be computed easily from the measured data because the usual measurement device only measures the surface EMG which is an weighted sum of EMG signals from muscles close to the sensor. Therefore it is rather difficult to estimate conditional probabilities from the data. One possible approach is to regard those nodes corresponding to the muscles as hidden nodes and try to estimate those conditional probabilities by using standard techniques such as EM algorithm or MCMC algorithm.

3.2 Representation by model transition Bayesian network

Since a node and an arc represent a model which is true at a certain time and a wolrd change respectively in a model transition Bayesian network, each node in the network corresponds to a possible world model in modal logic. Moreover, the structure of a model transition Bayesian network is analogous to that of HMM since each node represents a state and an arc represents a state transition. The advantage of this formalism is that it can represent a set of facts which are true at the same time in a single node. For example, simultaneous activations of a pair of antagonizing muscles can be naturally represented by a node which contains both of an active flexor muscle and an active extensor muscle. Another example is representation of different usage of arm segments; a mode using the entire arm from an upper arm to a wrist as a single unit, and another mode using their parts collaboratively to achieve elastic motion. These two modes can be represented by different sets of literals of different angular velocities and accelerations of each joint. Note that time sequence can be represented by a transition from a node to another node.

Figure2 shows an example of this representation: Figure2(a) shows how to represent a node, and Figure2(b) shows state transitions of "reflection" and "the release of antagonizing muscle usage".

One fundamental issue of this approach is how to define the original joint probability distribution of events. Since each node represents a model which becomes true in a certain time, we can define joint distribution by assigning time stamp for each literal in each node. Another fundamental issue is its representation complexity. Since each node represents a model, the representation of an entire network suffers from so called "frame problem". To avoid the problem, we need to introduce a computation schema such as Situation Calculus or Event Calculus. Since this formalism has very rich expressive power, we need rather complex inference mechanism.

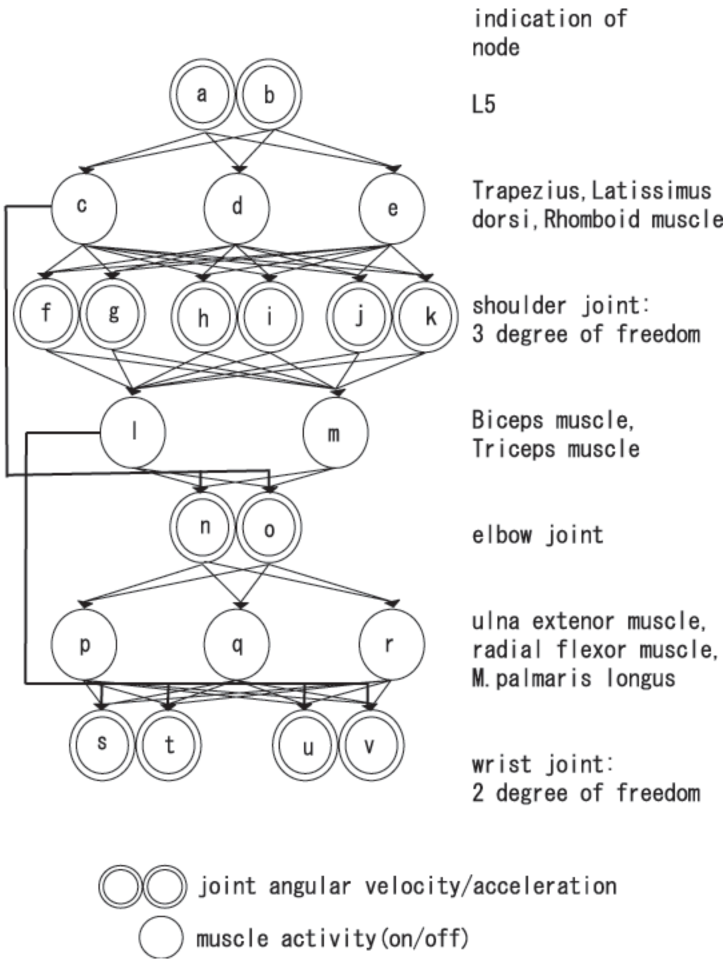


Figure 1: Musculoskeletal Bayesian network

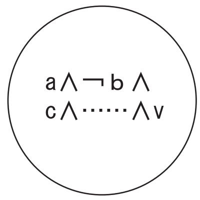
4 Assignment of Conditional Probabilities based on Observasional Data

We have several sensors for measuring cello playing: a motion capturing system, an EMG sensor, and a pressure sensor attached to the instrument. Among them, a motion capturing system provides data for computing coordinates, angular velocity and acceleration of each joint axis by properly attaching markers reflecting infrared. Actually we need proper preprocessing such as filtering and segmentation. Currently we are developing a new algorithm for segmentation and it will be reported elsewhere. Since the EMG data are not reliable because of the indirect measurement from surface and its inability to measure absolute EMG values, we cannot compute the conditional probabilities of both joints and muscles correctly. One possible approximation is to apply rough discretization and convert the original continuous data to categorical data. Simplest categorization is to classify the data into two classes; active or inactive. Another possibility is to model each group of related muscles by another Bayesian network and estimate activity of each muscle separately. In the latter case, we need to obtain data for each muscle activated independently.

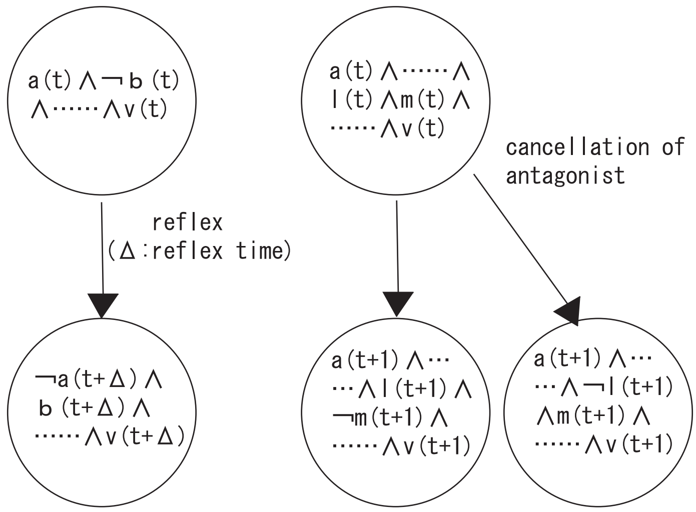
Our final goal is to compute such probabilities as Pr(antagonizing-muscles-on | wrist-snap-on), or Pr(antagonizing-muscles-on). The development of such computation algorithms are our future work.

5 Conclusion and Future Work

In this paper, we presented two possible representation schemas for expressing performance skill in terms of Bayesian network. The first one is a musculoskeletal Bayesian network which represents the musculoskeletal system naturally. It models rotation of each joint axis and the affect of related muscles. The second is a model transition Bayesian network which represents state transitions of possible activation patterns of muscle usage as



(a)node representation example



(b)transition example

Figure 2: Model transition Bayesian network

well as joint usage patterns.

The next work to be done is to establish algorithms to compute conditional probabilities of nodes and verify useful hypotheses by computing various joint/conditional probabilities. To achieve this goal, we need to find an algorithm to compute necessary parameters using available data and by applying such algorithms as EM algorithm or MCMC algorithm.

More concretely, we want to model a complex movement of the upper body including the spine, the upper arm, the fore arm, the wrist and fingers by both professional and amateur players. Note that down-bowing and up-bowing have different optimal musculoskeletal positions. It is true that they are both lead by the wrist movement. However, the direction of the wrist movements are different from each other; in case of down-bowing, the little finger side of the wrist leads the entire movement, whereas in case of up-bowing, the forefinger side leads it. The result is that the angles between the hand and the forearm are different from each other. This change of taking the optimal configuration should be attained instantaneously when one changes the bowing direction. More precisely, the configuration change should be achieved just before the bowing direction change. What cellists do for this purpose is to roll their wrists. Physically, one achieves the rolling of the wrist by utilizing inertia force of the hand movement. Biomechanically, one needs to use antagonizing muscles of the upper arm to make the wrist joing free. Neurophysiologically, it is considered that this instantaneous change can only be achieved by spinal reflex. It is one of the most important future research for us to reveal the true mechanism behind the movement stated above.

Furthermore, we recognized the importance of relating Bayesian network and qualitative reasoning during the process of building our fundamental scheme of tacit knowledge modelling. Although the combination seems promising, we could not achieve to provide a precise combination scheme. This is another important future work.

The use of still more expressive formalism of probabilistic logic programming such as PRISM or SLP is also interesting future research to be conducted.

Neurophysiologically, as mentioned above, spinal reflex seems to play a very important role in achieving skillful performance. It is our final goal to provide a convincing explanation on such a deep mechanism related to neurophysiology.

References

[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.

[Igarashi 02] Igarashi.S., Ozaki.T, and Furukawa.K.. "Respirarion Reflecting Musical Expression: Analysis of respiration during musical performance by Inductive Logic Programming." Proc. of the Second International Conference on Music and Artificial Intelligence. 2002.

[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)

[Muggleton 96] Muggleton.S.."Stochastic logic programs." In L. de Raedt, editor, Advances in Inductive Logic Programming. pp.254-264. IOS Press. 1996.

[Nakagawa 02] Nakagawa.S.. "Beyond HMM and Trigram for Automatic Speech Recognition." Journal of the Japanese Society of Artificial Intelligence. Vol.17 No.1 pp.35-40. 2002. (in Japanese)

[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.

[Shibuya 94] Shibuya.K, Sugano.S, and Kato.I.. "Skill-analysis of the bowing motion in violin playing." The Japanese Journal of Ergonomics. Vol.30 No.6 pp.395-403. 1994. (in Japanese)

[Russell 95] Russell.S. and Norvig.P.. "Artificial Intelligence, A Modern Approach" Prentice Hall. 1995.

[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.