IPSJ Transactions on Mathematical Modeling and Its Applications Vol. 4 No. 3 49–58 (July 2011)

Regular Paper

Information-theoretic Analysis for Understanding the Behavior of Song Learning by the Bengalese Finch

Khan Md. Mahfuzus Salam,[†1] Tetsuro Nishino,[†1]
Kazutoshi Sasahara,[†2] Miki Takahasi[†2]
and Kazuo Okanoya[†2]

Songbirds have been actively studied for their complex brain mechanism of sensor-motor integration during song learning. Male Bengalese finches learn singing by imitating external models to produce songs. In general, birdsong which is string of sounds is represented by a sequence of letters called song notes. In this study, we focus on information-theoretic analysis of these sequential data to explore the complexity and diversity of birdsong, and learning process throughout song development. We design and develop the analysis tool which has many features to do analysis for the sequential data. For experiment, we employ thirteen male Bengalese finches, each with different bouts of song data. By applying ethological data mining to these data, we discover that the finches follow two types of song learning mechanism: practice mode and adopt mode. In addition, over the analysis we find that it is possible to visualize the song features, e.g., traditional transmission, by contour surface diagram of the transition matrix. Furthermore, we can easily identify the families from these contour surface diagrams, which is a very challenging task in general. Our obtained results indicate that analysis based on data mining is a versatile technique to explore new aspects related to behavioral science.

1. Introduction

Ethology is the scientific study of animal behavior for exploring mechanisms underlying diverse forms of behaviors, from unlearned stereotyped ones to learned flexible ones. Songbirds have been actively studied as a good ethological model for their complex brain mechanism of sensor-motor integration in song learning. The Bengalese finch (*Lonchura striata var. domestica*) is a domesticated strain of a Southeast Asian finch, the white-rumped munia (*Lonchura striata*) and, it has been a popular subject for neurobiological and ethological studies on birdsongs for its unique song features. Birdsongs are strings of sounds represented by a sequence of letters known as a song note. The males of this species acquire their own songs by learning from external tutors (father or other males) during a specific period, between the nestling and fledgling stages. The learned features of the songs change very little during the matured stage of life. Their songs are used for courtship display. Two types of song features are preferred by female birds: performance- and elaboration-related traits. The performance-related trait is associated with the extent of song production, song rate, song duration, song speed, song amplitude, etc. In contrast, the elaboration-related features are associated with song rules and complexity.

Some previous study on the performance-related traits reported that, female preference is positively correlated with song duration and number of note types, which is negatively correlated with peak amplitude frequency of the song [14]. The study on elaboration-related traits reported that, female birds prefer ‘Sexy notes’ that have a complex song structure [14]. Besides, Chatfield and Lemon mentioned the importance of information-theoretic measures for the analysis in the field of animal behavior study [7].

Sasahara, et al. reported 2 types of development of song syntax [15]. Their analysis result was based on number of edges that is required to represent the song syntax by an automaton. Such interesting phenomena during song development motivated us to study the learning process of Bengalese finch song. No previous research was reported to understand the process by which songbirds learn to sing by doing analysis from information-theoretic viewpoint, employing the developmental song data. In the present study, we focused on elaboration-related features from 2 aspects: the first is analysis of song development and the other is comparison between the songs of the parent and progeny to understand the learning. Our aim is to conduct information-theoretic analysis on sequential song data to explore the diversity of learning birdsong throughout the song development. For this purpose, we design and develop the analysis tool. Using this tool our experiment on Bengalese finches shows, in practice mode, some finches sing complex songs in the early stages of development, and gradually crystallized the songs by eliminating extra transitions. On the other hand, some other finches do

†1 Information and Communication Engineering, The University of Electro-Communications, Tokyo, Japan
†2 Laboratory for Biolinguistics, RIKEN Brain Science Institute (BSI), Saitama, Japan

not apply practice mode. Their song production counts toward selecting, constructing, and maintaining behavioral outcomes, hereby it is called adopt mode.

2. Preliminaries

This section briefly introduces the theoretical foundations of birdsong and its representations, information-theoretic measures, and song model representation by k-reversible language.

2.1 Language and Birdsong

Humans use language to express emotions or to communicate with other humans, and language is considered unique to the human race. However, other living creatures, too, communicate vocally, for example, songbirds and dolphins, which have complex vocal communication homologous to human language.

In the 1960s, the prominent linguist C.F. Hockett divided "the design features of language" into 13 [10]. The proposed design features are as follows: (1) auditory-vocal channel, (2) broadcast transmission and directional reception in auditory signals, (3) rapid fading of auditory signals, (4) interchangeability in communication, (5) total feedback, (6) specialization, (7) semantics, (8) arbitrariness, (9) discreteness, (10) displacement, (11) productivity (creativity), (12) traditional transmission, and (13) duality. On comparing animal and human communication, Hockett concluded that of the 13 features, only 2 — "traditional transmission" and "duality" — were not observed in any animals. Here, "traditional transmission" indicates that linguistic knowledge is passed on from one generation to the next through learning, and "duality" indicates that particular sound elements have no intrinsic meaning but combine to form structures (e.g., words and phrases) that have meaning. However, recent ethological studies have revealed that traditional transmission is found in the vocal communication of songbirds and whales, and these organisms can combine a few discrete sound elements and so even exhibit duality in communication [5]. Therefore, these properties are not unique to human communication.

The song of the Bengalese finch has a more complex structure than that of other songbirds, such as zebra finches (*Taeniopygia guttata*) [3]. According to recent studies, the courtship songs of the Bengalese finch have unique features and are similar to human language [5]. Some research shows that the language model of this bird can be represented by a k-reversible automaton [12]. Thus, because of the structural and functional similarities in vocal leaning between songbirds and humans, the former have been actively studied as good linguistic models. In particular, the song syntax of the Bengalese finch sheds light on the biological foundations of syntax in humans [1].

2.2 Song of the Bengalese Finch and Its Representation

Recent studies on Bengalese finches have shown that the songs of the male birds are neither monotonous nor random; they consist of chunks, each of which is a fixed sequence of a few song notes. The song of each individual can be represented by a finite automaton, which is called song syntax (**Fig. 1**) [5]. Thus, the songs of Bengalese finches have "double articulation," which is one of the important structures of human language (i.e., a sentence consists of words, and a word consist of phonemes). Song syntax is controlled by song control nuclei in the brain. The hierarchy of song control nuclei directly corresponds to the song hierarchy [4].

Bird song analysis requires song data that have been recorded in a suitable environment. From the recorded wave data, spectrograms are obtained, and these are used as the standard representation of the song. Below are brief explanations of some general terms used in birdsong research.

Song note: A song note is a symbol assigned for an independent pattern that appears in a sonogram as seen in **Fig. 2**. It is also referred to as a song element or behavioral element. From the definition, we can say the text data comprising symbols (such as *a, b,* and *c*) are called song note sequences. Song notes are analogous to phonemes in human language.

Chunk: A fixed sequence of song notes is called a chunk. In Fig. 2, for example, the chunks are *ab, cde,* and *fg*. Chunks are analogous to words in human language.

Song unit: A song unit consists of many chunks. Song units are analogous to sentences in human language.

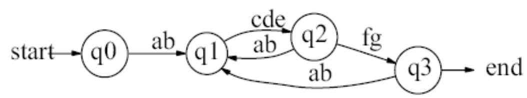

Fig. 1 Courtship song syntax represented by an automaton.

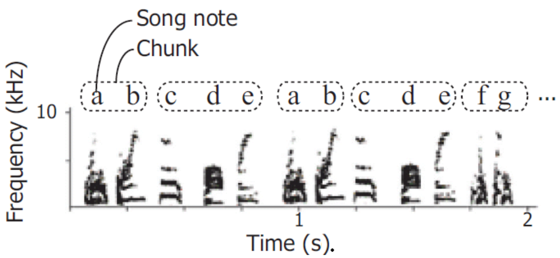

Fig. 2 Grayscale spectrogram of a Bengalese finch's song.

Song bout: A song bout consists of one or more song unit that is a continuous song produced by the bird in one time. Song bouts are analogous to paragraphs in human language.

2.3 Information-theoretic Measures

In this section, we briefly discuss the information-theoretic measures that describe the features in behavioral strings and help understand their diversity.

Transition matrix: In birdsong research, a transition matrix is widely used to understand syntactical complexity. A transition matrix is one that shows note-to-note transition information. A transition probability matrix can be obtained by dividing the note-to-note outcome by the total number of transitions. This is the most common and important way to represent transition information, and other properties can be analyzed using this matrix.

Linearity: Scharff and Nottebohm introduced the measure of linearity for estimating the ordering complexity of notes [9]. The linearity index score is calculated from the number of note types and transition types as follows:

$$S_{Linearity} = \frac{number\ of\ different\ notes/song}{number\ of\ transition\ types/song}$$

The linearity index score provides an estimate to predict the next note in a song when the previous note is known. In a completely linear song sequence, each note has only 1 transition type. Thus, a complete linear song has a linear index score of 1. If some notes have more transition types, the score becomes less than 1. Therefore, a lower linearity index score reflects a more syntactically complex song.

Entropy: Shannon introduced the concept of information entropy and discussed the details of entropy in printed English [8]. Information entropy is a statistical measure of the uncertainty associated with a random variable. It quantifies, in certain sense, how much of information is produced on average for each letter of a certain text. Information entropy is denoted by

$$H(x) = -\sum_{i=1}^{n} P(x_i) \log P(x_i)$$

where, n is the elements and $P(x_i)$ is the probability of appearing symbol x_i in the sequence.

Chatfield and Lemon devised a method to calculate higher-order entropies based on the Markov model for sequential data [6),7)], which is defined by

$$H_n = H(n) - H_{n-1} \text{ where, } n > 1$$
$$H(n) = -\sum_{pre,cur} P(X_{pre}, X_{cur}) \log P(X_{pre}, X_{cur})$$

where, n is the order, $P(x, y)$ is the probability of appearing symbol x after symbol y. However, this Markov model estimation with improperly higher-order entropies probably does not consider all the transitions that the true source produces or detects them with an incorrect frequency. Therefore, it produces a deceptively low and inaccurate entropy estimate.

n-gram statistics: An n-gram is a sub-sequence of n items from a given sequence. The items can be phonemes, syllables, letters, or words depending on the application. The size of the ordered list of elements is denoted by n. An n-gram of size $n = 1$ is referred to as a "unigram"; $n = 2$, as a "bigram"; and $n = 3$, as a "trigram." One with $n = 4$ or more is simply called an "n-gram". "$(n - 1)$-order Markov models" are language models built from n-grams. In ethology, n-gram statistics is used to understand frequency distribution and the hierarchy of behavioral patterns.

3. Data Mining and Information Extraction

This section describes the application of the data mining technique to extract information from behavioral sequences.

3.1 Data Mining

Data mining is the process of analyzing data from different perspectives and summarizing it into useful information. Various approaches can be used to analyze data for data mining. One approach is to use an analyzing tool that allows users to analyze data from many different dimensions, categorize it, and summarize the relationships identified. In general, data mining is the process of finding correlations or patterns among relational datasets.

3.2 Mining Behavioral Sequential Data

In general, animal behavior is recorded as sequential data of behavioral events. The same symbol is assigned to an identical behavioral event type, and the behavioral data is converted into text data, which are then used for data mining. Such sequential data of animal behavior can be analyzed for different statistical and information-theoretic measures, such as the transition matrix and the first-order Markov chain, and entropy [12]. Special tools are required for dealing with complex behavioral data, and these should be used together with conventional tools. Data mining is one way to do handle complex behavioral data. The current data mining process has many well-established techniques for pattern extraction, clustering, modeling, etc. [13]. It can help find units from behavioral sequences and extract the rule that governs them. For obtaining significant information from animal behavior, we have to carefully select elemental data mining techniques suitable for behavioral sequences, which might be different from both word sequences in natural languages and biological sequences like DNA, and we have to then use them with proper modification in the context of ethology.

In ethological studies, there are 3 steps for understanding animal behavior. First, a behavioral phenomenon is observed and recorded; second, on the basis of the recordings, a hypothesis is formulated to explain the behavior; and third, experiments are designed, performed, and evaluated to test the hypothesis [13]. If these procedures are followed correctly, better predictions can be made concerning animal behavior, which in turn provide insights into human behavior. However, developing a hypothesis on the basis of recordings is not always easy if the data is vast or complex. Therefore, we apply data mining technology in ethology. In particular, we studied the application of ethological data mining from sequential animal behavior. The following (**Fig. 3**) shows the process flow we used.

In this study, we employ the song of the Bengalese finch to analyze song data

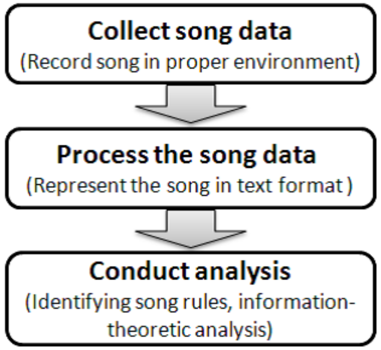

Fig. 3 Process flow of an ethological study on birdsong.

during song development. To collect the song data, all birds were raised in the same environment inside a cage in laboratory. Their vocalizations were recorded to analyze their songs. Each bird was individually placed in a soundproof room, and its vocal output was recorded with a directional microphone and a DAT recorder.

The recorded songs were analyzed with sound analysis software to generate a spectrogram, which was used to convert sounds (WAV format) to texts. For the song of the Bengalese finch, we obtained a spectrogram where the different song notes were separated by a considerable gap. Each identical pattern was labeled with a similar symbol like *'a,' 'b,'* etc. Thus, the song was represented in text format.

This analysis was performed manually on the basis of the phonological properties of song notes. The text data of the song were organized by bouts, which is the unit of time for which songbirds sing at a stretch. The following are examples of simple bout data.

(Bout 1) abcdbefggabcdbefggabcdbefggabcdbefgghijkli bkmggabcdbefggabcdbefg...

(Bout 2) abcdbefggabcdbefgghijklibkmgggabcdbefggab cdbefgghijklibkmgggabcdb...

If we carefully look at the song data sequence we can easily find some pattern in the sequence. When the bouts are delimited into segments at 'a,' they are found to consist of 2 types of song units — *"gabcdbefgg"* and *"gabcdbefgghijjklibkmggg"* — which repeatedly appear in other bouts. Thus, by some pre-processing, we converted the song data into text for further analysis. For data mining from the behavioral sequence, we developed a tool called *EUREKA*, which stands for **e**thoinfomatical **u**tilities for **r**ule **e**xtraction and **k**nowledge **a**cquisition [16]. EUREKA is a utility suite used for the following analyses:

- Information-theoretic analysis
- Extraction of probabilistic behavioral rule (n-gram model)
- Extraction of deterministic behavioral rule (deterministic finite automaton)

We design the features of the *"StringStat"* module of EUREKA tool and develop the functionality of different information-theoretic measures to meet the requirement of detail analysis. The *"StringStat"* module deals with the analysis of information-theoretic measures and enables the following translation in the context of ethology:

Linearity: Analysis of behavioral diversity.

Entropy: Analysis for behavioral uncertainty.

Transition Matrix: Representation of the transition probability of behavioral event types.

n-gram Statistics: Analysis of different n values in terms of frequency distribution and hierarchy of behavioral patterns.

4. Data Mining from Birdsong

In this section, we present our results from 2 aspects. The first is the analysis of song development and the other is comparison of song property between the parent and progeny. We consider some of the measures in our analysis provided by "StringStat".

4.1 Description of Data

In the current study, we examined 9 juvenile birds to study song evolution and 4 parent-progeny pairs to compare the songs between parent and progeny. It has been reported that young male birds learn songs from their fathers in the first 120–130 days after hatching [2]. The principal learning period is considered the age of 60–130 days. The sounds that are made during the first 60 days after hatching are distorted and do not considered as songs.

Table 1 shows the days when the data was recorded for each individual that is used for analysis in the present study. Song data was recorded with intervals but was not fixed for different birds. For the convenience of analysis, we categorized those into 5 different stages.

Table 1 Young birds and their age (in days) when song was recorded.

Bird name	Age in days				
Stage	1st	2nd	3rd	4th	5th
Lao	61	88	95	102	116
RAo	61	76	87	102	127
LDai	73	80	94	108	129
Shiro	64	78	92	100	120
RMo	70	77	91	99	125
RKi	70	79	93	106	127
LShiro	73	85	93	100	121
Atama	62	69	89	99	124
LMizuiro	74	81	94	108	130

Table 2 shows the parent and progeny relationship for the birds:

Table 2 Relationship between parent and progeny.

Parent	Progeny
Sankakukoshitya	Lao
Bakatono	RAo
Katsuo	LDai
Kuroshiro	Shiro

4.2 Common Parent-progeny Findings

We present our results regarding features that were common to the songs of the parents and progeny. For this analysis, we considered only the transition matrix and transition types. We also present a new technique for visual representation of the song on the basis of the transition matrix.

The birds have 7–14 different note types. Our findings show that every bird has 1 or 2 dominant notes, while other notes are used to generate variations in the song. **Figure 4** shows the graph of rank vs. frequency for the different pattern types from the song sung by the bird *Shiro* at age 120 days. X-axis corresponds to the rank for n-gram pattern types where the value of n is from 1 to 10, and Y-axis corresponds to frequency. The complete song consists of 10 bouts comprising 816 song notes. The song note types are labeled as *a, b, c, d, e, f, g,* and *h*. We found that of these song notes, d appeared 317 times, and f appeared 228 times. Further, a appeared 123 times, and the other notes

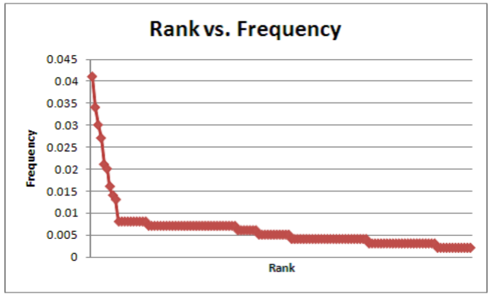

Fig. 4 Rank vs. frequency of n-gram types.

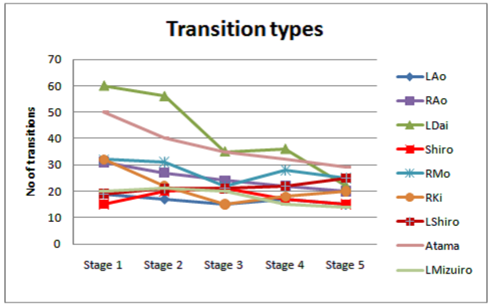

Fig. 5 Change in number of transition types during song development.

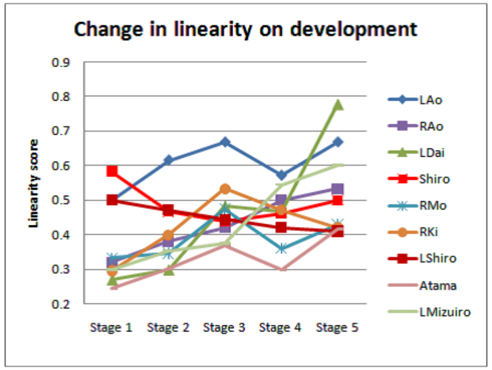

Fig. 6 Change in the linearity index score during song development.

appeared less than 100 times. When we increased the value of n, we found that *dd* appeared in the song sequence 264 times; *ff*, 159 times; *ddd*, 211 times; *fff*, 102 times; and *dddd*, 158 times. From this simple analysis, we found that the notes d and f are the dominant song notes for *Shiro*.

4.3 Analysis on Evolution of Song

Here, we present the results of evolution analysis. We considered only the young bird's song data for this analysis. Although *"StringStat"* module provides different information-theoretic measures, we focused on linearity, entropy, and n-gram statistics in particular transition types.

In the early stage, most birds have a relatively large number of transition types in their songs. Eventually, the noise transitions are reduced, and the birds produce relatively small patterns. However, some birds produce songs with a relatively small pattern from an early age, and the pattern does not change later during song development. We found the song development of *Shiro* and *LShiro* to show such a trend (**Fig. 5**).

The linearity index score indicates the complexity of the song. Our investigation of this measure showed that all the birds in this study sang complex songs in the beginning, although after the development period, the crystallized songs were syntactically simpler. Again, the experimental results showed that *Shiro* and *LShiro* started singing less complex songs from the beginning, and the linearity index score of these birds changed only slightly during song development (**Fig. 6**).

By definition, the linearity index score and entropy value are closely linked because both measures reflect the syntactical complexity of note-to-note transition. The following graph shows the correlation between the linearity and entropy value for all 9 progenies in their song development period (days in Table 1). In our experiment, we also find a linear correlation between these 2 variables (**Fig. 7**) during song learning.

Based on the above three analysis our investigation indicates that we can divide the birds into two groups. One group with majority of birds in their early stages of development has high transition rate and low linearity index score while producing the song. Those complex songs are gradually crystallized by the elimination of extra transitions. We call this learning process as "practice mode".

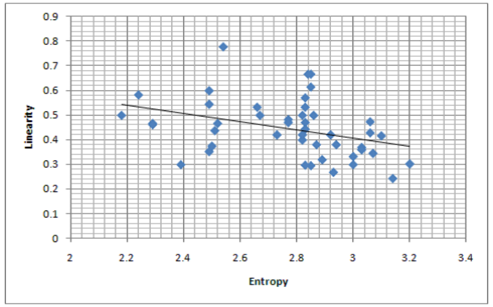

Fig. 7 Relationship between linearity and entropy during song development.

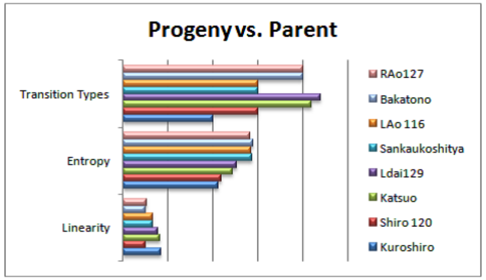

Fig. 8 Comparison of different measures between parent and progeny.

On the other hand, other group with minority birds, the rate of changing values for the number of transition and linearity index score is very low. During song production all activity counts toward selecting, constructing, and maintaining behavioral outcomes. This implies that the fundamental, self-generated activity plays an important role in the development of behavioral functions, both perception and cognition related. We call this learning process as "adopt mode".

4.4 Parent-progeny Comparison of Songs

This section shows results of a comparison between the songs of the parent and progeny. For this analysis, the song data of both the parent and progeny were considered. We focused on different information-theoretic measures such as linearity, entropy, and n-gram statistics in particular transition types.

Figure 8 shows the comparison of information-theoretic measures between parent and progeny. Four parents song data and 4 progenies matured periods song data (i.e., stage 5 in Table 1) has been employed here. It indicates that during "practice mode" learning, different measures converge to the value of the parents' song. Except the pair *"Kuroshiro–Shiro,"* all other pairs showed almost similar values of all measures between the parent and progeny. The learning process of *Shiro* was found to be "adopt mode". We can see from Fig. 8 that the transition types and linearity index score of this bird was very different from those of his father; in fact, *Shiro's* song was more complex than his fathers'.

4.5 Visualization of Song Features

In this section, we present a new technique for visual representation of the song on the basis of the transition matrix. To explain the usability of this visualization technique we also show contour surface diagrams of different bird families.

A contour surface diagram can be used for visual representation of the song property. In general, surface charts are useful 3D chart types; they have 3 true data dimensions and can illustrate data reasonably well. Those charts are useful to show how a variable (Z) changes according to 2 other variables (X and Y). Contour graph is a kind of surface chart containing regions colored according to the Z value. Essentially, they are 2D top views of 3D surface charts.

In this paper, to visualize the bigram property based on the transition matrix, we use the contour graph in an unconventional fashion. Where, the symbols (such as a, b, and c) are arranged on X and Y axes in the alphabetical order to generate a graph. We found that using contour graph is a good technique for visual representation of the bigrams. **Figure 9** shows a contour graph for the birdsong produced by bird *RAo* at age 127 days.

Figure 10 shows the contour diagrams of the transition matrix for the songs of the parents and progeny. The contour diagrams for the song at the early age and matured age are shown for the young birds. From Fig. 10, we can easily visualize 2 properties of the songs of the Bengalese finch: (1) Contour diagrams can visually display the unique song features of a particular bird family. We can easily distinguish different bird families from the contour diagrams. (2) Second, although there are differences in song properties at the early and matured stages, the major features are present at the early stage, and as the noise transitions reduce, the songs of the young birds eventually converge in to their fathers'

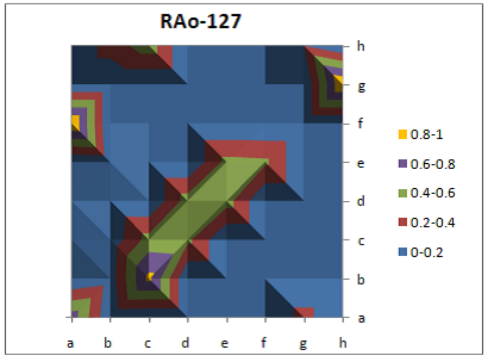

Fig. 9 Contour surfce diagram of bird RAo.

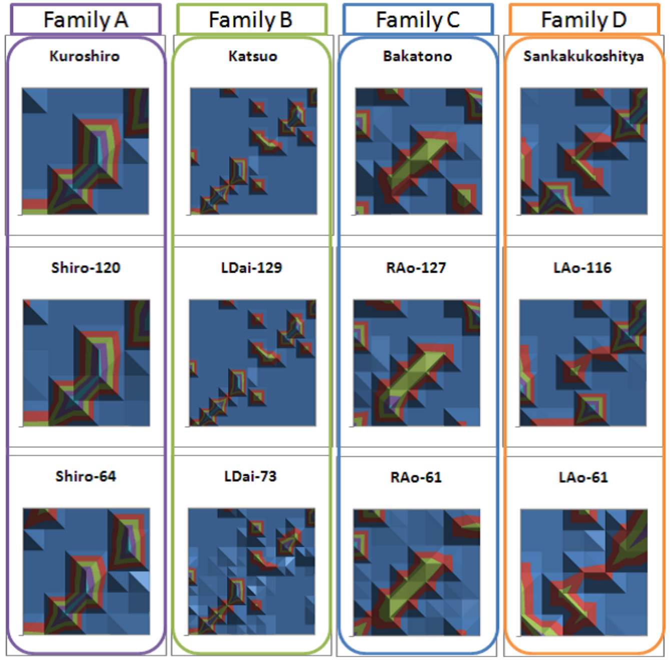

Fig. 10 Contour diagrams of different families.

songs. That is a clear evidence of "traditional transmission" feature mentioned by Hockett.

There are 3 main purpose of proposing this visual representation: (1) since the transition matrix shows only numbers, it is difficult to understand the song patterns from the matrix, but if we represent the song as a contour surface diagram, we can easily visualize its transition properties. (2) The corresponding symbols of father and progeny are directly related. Similar patterns of father and progeny is assigned with a same symbol. For that reason, even if one pattern is changed with a different symbol, obtained contour diagram can be different but will be similar for the particular family. (3) If we want to identify the families based on the properties of their song it will be a very difficult task for making an algorithm by applying some clustering technique. But from the contour diagrams we easily identify the families which make the task simple.

5. Conclusions

This paper reported on information-theoretic analysis of the sequential data of Bengalese finch to explore the learning process throughout song development. We show the effective use of the data mining for the birdsong research, in general, in the research of animal behavior. By applying ethological data mining to the birdsong data, we discover that the finches follow two types of song learning mechanism: practice mode and adopt mode, which is a new finding related to learning mechanism of birdsongs. In practice mode, some finches sing complex songs in the early stages of development, and gradually crystallized the songs by eliminating extra transitions. On the other hand, some other finches do not apply practice mode. Their song production counts toward selecting, constructing, and maintaining behavioral outcomes, hereby it is called adopt mode. Thus, such analysis provides scope for closer examination of a large amount of data, whereby useful information can be extracted from them.

In addition, we showed a new technique to visualize the features of behavioral sequences. Over the analysis we find that it is possible to visualize the song features, e.g., traditional transmission, by contour surface diagram of the transition matrix.

Our obtained results indicate that analysis based on data mining is a versatile technique to explore new aspects related to behavioral science. By applying the findings of the present study, we will be able to analyze animal behavior more precisely. For example, our analysis shows that kuroshiro — father of Shiro, is the most stupid bird in terms of song complexity. We may make a hypothesis that, if the father is stupid than son become smarter and that can be the reason of having adopt-mode learning process. Further study is necessary to establish this hypothesis. Beside this, we could obtain a better understanding of the features of dominant song notes to compare their sound properties and effect on female preference. Thus, the findings in this current study shed lights on new aspects of future research related to behavioral science.

References

1) Doupe, J. and Kuhl, P.K.: Birdsong and Human Speech: Common Themes and Mechanisms, *Annual Reviews Neuroscience*, Vol.22, pp.567–631 (1999).
2) Catchpole, C.K. and Slater, P.J.B.: *Bird Song: Biological Themes and Variations*, 2nd edition, Cambridge University Press (2003).
3) Honda, E. and Okanoya, K.: Acoustical and Syntactical Comparisons between Songs of the White-backed Munia (Lonchura striata) and Its Domesticated Strain, the Bengalese Finch (Lonchura striata var. domestica), *Zoological Science*, Vol.16, pp.319–326 (1999).
4) Nishikawa, J. and Okanoya, K.: Dynamical Neural Representation of Song Syntax in Bengalese Finch: A Model Study, *Ornithological Science*, Vol.5, No.1, pp.95–103 (2006).
5) Okanoya, K.: Song Syntax in Bengalese Finches: Proximate and Ultimate Analyses, *Advances in the Study of Behavior*, Vol.34, pp.297–346 (2004).
6) Suzuki, R., Buck, J.R. and Tyack, P.L.: Information Entropy of Humpback Whale Songs, *Journal of the Acoustical Society of America*, Vol.119, pp.1849–1866 (2006).
7) Chatfield, C. and Lemon, R.E.: Analyzing Sequences of Behavioral Events, *Journal of Theoretical Biology*, Vol.29, pp.427–445 (1970).
8) Shannon, C.E.: Prediction and Entropy of Printed English, *Bell System Technical Journal*, Vol.3, pp.50–64 (1950).
9) Scharff, C. and Nottebohm, F.: A Comparative Study of the Behavioral Deficits following Lesions of Various Parts of the Zebra Finch Song System: Implications for Vocal Learning, *Journal of Neuroscience*, Vol.11, No.9, pp.2898–2913 (Sep. 1991).
10) Hocket, C.F.: The Origin of Speech, *Sci. Am.*, Vol.203, p.88 (1960).
11) Thelen, E. and Smith, L.B.: *A Dynamic Systems Approach to the Development of Cognition and Action*, The MIT Press (1994).
12) Kakishita, Y., Sasahara, K., Nishino, T., Takahasi, M. and Okanoya, K.: Ethological Data Mining: An Automata-Based Approach to Extract Behavioral Units and Rules, *Data Mining and Knowledge Discovery*, Vol.18, No.3, pp.446–471 (2009).
13) Han, J., Cheng, H., Xin, D. and Yan, X.: Frequent Pattern Mining: Current Status and Future Directions, *Data Mining and Knowledge Discovery*, Vol.15, No.1, pp.55–86 (2007).
14) Soma, M., Takahasi, M., Hasegawa, T. and Okanoya, K.: Trade-offs and correlations among multiple song features in Bengalese Finch, *Ornithological Science*, Vol.5, No.1, pp.77–84, (2006).
15) Sasahara, K., Kakishita, Y., Nishino, T., Takahasi, M. and Okanoya, K.: Constructing Song Syntax by Automata Induction, *ICGI–2006* (Sep. 2006).
16) TNLAB, UEC & OKANOYA LAB, RIKEN Japan, *EUREKA*. http://sites.google.com/site/eurekawiki (last accessed: November 1st, 2010)

(Received November 17, 2010)
(Revised January 14, 2011)
(Revised(2) February 19, 2011)
(Accepted March 17, 2011)

Khan Md. Mahfuzus Salam graduated M.Eng. from The University of Electro-Communications in 2009, and enrolled in Ph.D. course from 2009. Beside Birdsong analysis research, he has interest in Data mining, Artificial intelligence and Social Simulation on Agent based systems. He has been awarded scholarship from Hirose International Scholarship Foundation for his M.Eng. and Ph.D. study. He is a student member of IPSJ and Society For Neuroscience.

Tetsuro Nishino graduated from department of Mathematics, Waseda University and continued his research and obtained a D.Sc. in Mathematics in 1991. He was a researcher at Tokyo Research Laboratory, IBM Japan during 1984–1987. He was an Associate Professor at School of Information Science, Japan Advanced Institute of Science and Technology, Hokuriku during 1992–1994. In 1994, he joined at Department of Communications and Systems Engineering, The University of Electro-Communications as an Associate Professor and in 2006 he became a Professor. He received The Funai Information Technology Prize (2003) and IBM Faculty Award (2008).

Kazutoshi Sasahara graduated from Graduate School of Arts and Sciences, the University of Tokyo and obtained a Ph.D. in 2005. He continued his research at Laboratory for Biolinguistics, RIKEN Brain Science Institute, during 2005 and 2010 as a JSPS research fellow. His research major is Artificial Life and Biolinguistics. He is a member of The Physical Society of Japan, Information Processing Society of Japan, Society of Evolutionary Studies, Japan and Japan Ethological Society.

Miki Takahasi graduated from faculty of letters, Chiba University and obtained a Ph.D. in 2006. She continued her research at Laboratory for Biolinguistics, RIKEN Brain Science Institute, during 2005 and 2011. Since 2011, she is a JSPS research fellow at the University of Tokyo. She struggles with her 3 year-old daughter every day.

Kazuo Okanoya graduated from department of Psychology, Keio University and continued his research at University of Maryland, USA, and obtained a Ph.D. in Biopsychology in 1989. He was an assistant professor at department of cognitive and information sciences at Chiba University during 1994 and 2004. He then took the position of a laboratory head at Laboratory for Biolinguistics, RIKEN Brain Science Institute, during 2004 and 2011. Since 2011, he is a professor of psychology at the University of Tokyo. He is also a director of JST ERATO Okanoya Emotional Information Project. He enjoys Renaissance music and plays the lute.