生体信号のデコーディングによる神経情報処理の解析

Decoding of Biological Signal for Analyzing Neural Computation

船水 章大　　指導教員　高橋 宏知 講師

Akihiro FUNAMIZU　　(Lecturer Hirokazu TAKAHASHI)

Keywords: Reinforcement learning, Bayesian inference, Decoding, Auditory cortex, Dynamic brain

1. Introduction

Animals get substantial sensory evidences as inputs of brain, and execute actions as outputs to adapt environment. Here, I propose a hypothetical architecture of the transfer rules between brain and the inputs/outputs, as shown in Fig. 1. The transfer rules are categorized into multiple modules of three levels, i.e., algorithms, parameters and neural activities, and the weight of each module dynamically changes by a problem to solve and even during a problem. In contrast with my hypothesis, conventional studies usually merely focus on one of each level or only the neural activity, making it difficult to test the rules of highly adaptable brain. In my hypothesis, brain has a lot of algorithms and dynamically selects an algorithm to use. The algorithm consists of various parameters and a series of parameters to use also changes dynamically. The hardware, i.e., neural activity, for implementing the parameters also has various features such as spike counts and spike timings. To investigate the temporally dynamic brain, I test how the neural activities, parameters and algorithms changed dynamically.

2. Dynamic system of neural activities

2.1. Experimental objective

To test the dynamic system of neural activities, I focus on the tone frequency representation of auditory cortex in various tone intensity ranges. The auditory cortex is thought to have a distinct place code of tone frequency, or tonotopic map. However, the spatially focal activation as a function of test frequency is observed only at a low sound pressure level (SPL); at moderate or high SPLs, the activation patterns become less distinct across test frequencies, suggesting that the sound frequency is not represented by a simple place code of firing rates, but that the information is distributed spatio-temporally irrespective of the focal activation[1]. Here, I investigate the tone-frequency representation of auditory cortex in mid to high SPLs, and test how the frequency representation depends on tone intensities.

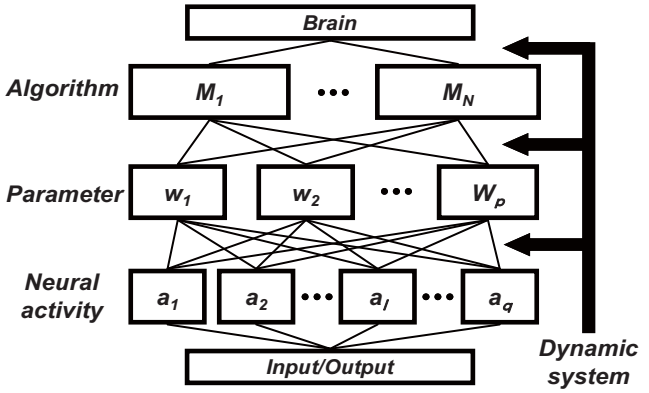

Fig. 1. Dynamic system of brain

2.2. Methods

All procedures were approved by my institutional committee and performed in accordance with "Guiding Principles for the Care and Use of Animals in the Field of Physiological Science" of the Japanese Physiological Society. Rats were anesthetized with isoflurane, and tone-evoked auditory cortical neuronal activities were recorded with a 24-probes microelectrode array. Test tones frequencies ranged between 5 and 50 kHz with 5 kHz increments and the intensities ranged between 50 and 70 dB SPL with 5 dB increments which corresponds to mid-intensity range. Each tone presented 20 times in pseudorandom order with inter-tone interval of 200 ms.

To identify spatio-temporal neural activity patterns important for the tone frequency representation, I conducted a recursive feature elimination (RFE). RFE extracted important spatio-temporal windows one by one without deteriorating a correct identification of a decoder. As the decoder of RFE, I employed support vector machine (SVM), k-nearest neighbor (KNN) and canonical discriminant analysis (CDA), and identified which decoder had the best correct identification. The correct identification was analyzed with 5-fold cross validation.

Correct identification of SVM with RFE was 68.5 ± 2.81% (mean ± standard errors, here and hereafter) in 12 subjects which was higher than that of KNN with RFE (59.36 ± 2.81%; Wilcoxon signed-rank test, $p = 0.0226$) and CDA with RFE (64.56 ± 2.46%; $p = 0.175$). Therefore, I employed SVM in the following analyses; I referred SVM with RFE as sequential dimensionality reduction (SDR).

With SDR-identified spatio-temporal neural activity patterns, i.e., SDR patterns, I first compared the correct identification with spatial patterns, in which temporal structures were eliminated, or high spike-rate patterns, in which windows with high spike rates were selectively extracted. I then measured a sparseness of SDR pattern, which I referred to as pattern sparseness. The pattern sparseness measures how sparsely population of neurons represent a given stimulus. Further, I investigated a spike rate and the weight parameters of SVM in each spatio-temporal window of SDR patterns. To test how the variety of spike rates across windows contributed to the classification, I first characterized the absolute weights and spike rates in each of the windows. I then focused on a window that had the highest weight in each SVM and characterized the normalized spike rate of the window with respect to a frequency difference to be classified (Δf). Since I had 45 SVMs a pairwise coupling, I

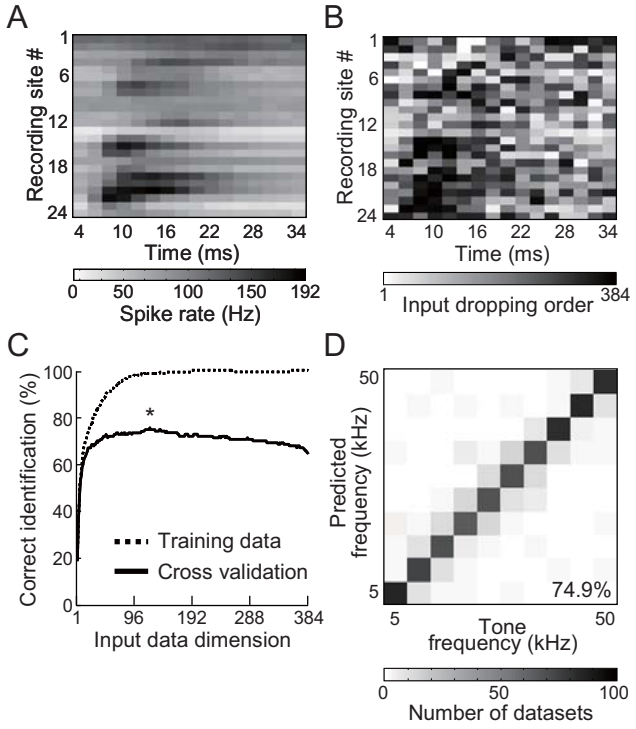

Fig. 2. Example of sequential dimensionality reduction

(A) Spatio-temporal patterns of spike rates grand-averaged across all test stimuli in a representative subject. (B) Dropping order of SDR. Dropping order of each window is shown in gray scale. Windows with small orders (white) dropped early. (C) Correct identification as a function of dimension of input data. Asterisk indicates the dimension with the highest correct identification. (D) Confusion matrix for cross validation. Vertical and horizontal axes indicate tone frequencies predicted and tone frequencies presented, respectively. The number of each prediction presentation pair is shown in gray scale. Correct identification in total is shown in the lower right. The input data dimension was 121, indicated by the asterisk in (C).

obtained 45 highest-weight windows for each subject.

2.3. Results

In a representative subject, Fig. 2A shows recorded auditory cortical neural activities which grand-averaged the spike rates for all test tones, and Fig. 2B shows the dropping order of SDR. SDR tended to extract high-spike-rate windows, but some low-spike rate windows were also extracted, implying an importance of low-spike-rate windows in frequency representation. Fig. 2C shows the correct identification of SDR. The highest correct identification was achieved during the dimensionality reduction, and all the test frequencies were succeeded to predict in SDR, as shown in Fig. 2D.

SDR achieved the best correct identification of 68.6 ± 2.76% at 93.8 ± 7.85 input data dimensions. Correct identification at 384 dimensions was 59.5 ± 2.91%, indicating that SDR identified approximately a quarter of spatio-temporal windows and significantly improved the correct identification (Wilcoxon signed-rank test, p = 4.88E-4). Moreover, SDR patterns led to better decoding than spatial patterns and high-spike-rate patterns, of which the correct identifications were 62.9 ± 2.50% and 61.7 ± 2.75%, respectively (Wilcoxon signed-rank test after Bonferroni

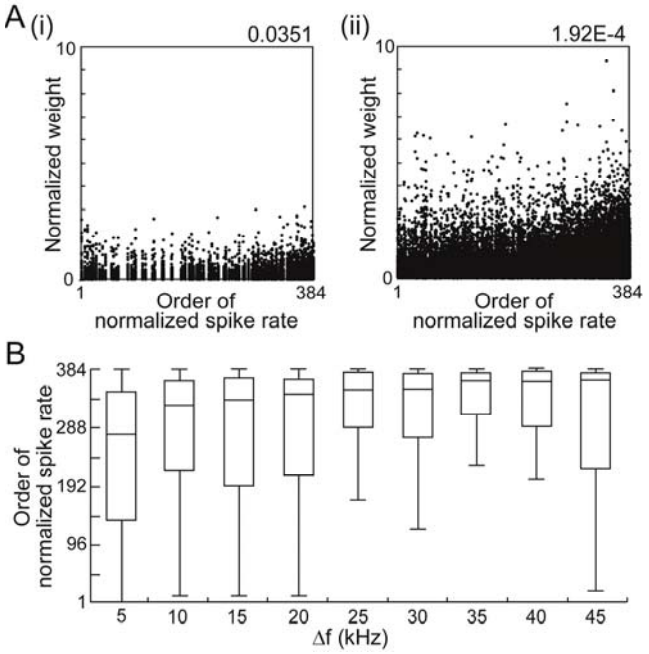

Fig. 3. Contribution of low-spike-rate windows

(A) Weight distribution of SVM as a function of rank order of normalized spike rate: (i), Representative subject; (ii), All subjects. Each window has 45 weights because there are 45 SVMs for a pairwise coupling, and thus, 45 weights are distributed at each rank order of individual subject. Correlation coefficients between the weight and rank order are shown in upper right. (B) Rank order of normalized spike rate in the highest weight window with respect to Δf. The spike rates depended on Δf (Kruskal-Wallis, p = 3.06E-7), and were positively correlated with Δf (r = 0.213, p = 5.52E-7)

correction, p = 1.47E-3 in both comparison). This result suggests that tone frequency is encoded in temporal as well as spatial structures of neural activities and low-spike-rate windows.

Pattern sparseness of SDR pattern decreased from the initial condition, i.e., 384 dimensions, by 0.0710 ± 0.0251, which was of significant reduction (two-side t-test, p = 0.0164). In addition, the correlation between dropping order of SDR and the normalized spike rate of each window were positive (r = 0.504). Taken together, SDR pattern tended to extract high-spike-rate windows and composed a highly dispersive pattern that probably offers an advantage of discrimination ability.

Investigation of SVM weights suggested that, in SDR patterns, the weights of SVM did not correlate with the spike rates as shown in Fig. 3A, suggesting that both high- and low-spike-rate windows are participated in tone-frequency representation. Further analysis of the weights showed that normalized spike rate of a window that had the highest weight in each SVM depended on Δf as shown in Fig. 3B, suggesting that low-spike-rate windows participated in fine frequency differentiation. This result suggests that the auditory cortex effectively overcomes the difficulty of local differentiation of frequencies in mid-intensity tones by exploiting low-spike-rate windows. Taking into account the frequency representations in low-intensity tones, in which tone frequencies are represented with high-spike-rate activities, neural features of tone-frequency representation in the

auditory cortex may dynamically change depending on sensory stimuli.

3. Dynamic system of parameters

3.1. Experimental objective

To test the dynamic system of parameters, I focus on the uncertainty and investigate the role in action selection. Neurophysiological studies show that the uncertainty is actually encoded in brain[2], leading to a hypothesis that brain uses the uncertainty of reward expectation, as well as the value, to guide behavior. However, studies of animal behaviors usually employed a standard reinforcement learning which ignores the uncertainty, making it difficult to investigate the role of uncertainty. In this experiment, I employ Bayesian Q-learning, where the uncertainties are implemented, and analyze rats' choice behaviors in a free choice task. I then compare the likelihood with standard Q-learning and investigate how the role of uncertainty changes during the task.

3.2. Methods

In the free choice task, rat selected either left or right hole, and received a reward of food pellet stochastically. Reward probability of each choice changed among six reward-probability states, i.e., 100%-66%, 66%-33%, 33%-0%, 66%-100%, 33%-66% and 0%-33% (reward probabilities of left-right choices, here and hereafter), when the frequency to select an optimal choice reaches 80% in last 20 trials.

Bayesian Q-learning assumes that the value of each choice has a distribution, served as the uncertainty of value, and the value is updated with Bayesian inference. I prepared three Bayesian Q-learning and four standard Q-learning models by testing all possible equations to update values with past action-reward sequences. I validated the likelihood of models with 2-fold cross validation. I also analyzed the likelihood in trials around state-change periods in which a reward-probability state changed another.

3.3. Results

I used two rats and analyzed 45 sessions of data. I first found that rats' choice behaviors were fit to Bayesian Q-learning which only tracks a positive reward information, i.e., RBQ-learning, as shown in Fig. 4A. Further analysis of likelihood showed that the choice behaviors matched the Bayesian Q-learning especially in low- and mid-reward-probability environments, while the behaviors were fit to a standard Q-learning in high-reward probability states as shown in Figs. 4B-E. Thus, the uncertainty is mainly important in low- and mid-reward-probability states, suggesting that the role of uncertainty for action selection dynamically changed by the context.

4. Dynamic system of algorithms

4.1. Experimental objective

For investigating the dynamic system of algorithms, I focus on the role of model-free and model-based algorithms, or strategies, in action learning. Recent studies suggest that model-free strategies, which learn action values without assuming a higher-order structure, are encoded in dorso-lateral striatum, and model-based strategies, which use a stored model of the task structure to guide choice, are encoded in ventro-medial prefrontal cortex[3]. Although, the studies predict the dual strategy system of brain, little is known about whether animals employ the two strategies dynamically depending on the context, or even during a

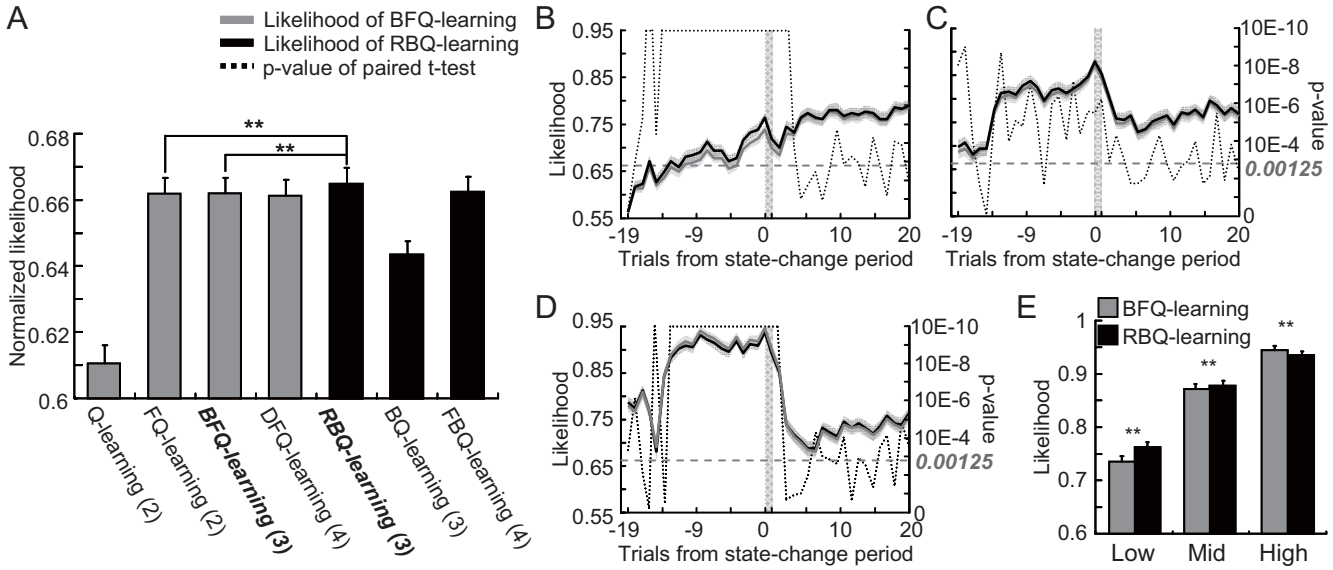

Fig. 4. Normalized likelihood of Bayesian Q-learning and standard Q-learning

(A) Result of 2-fold cross validation. Gray and black bars show standard Q-learning and Bayesian Q-learning, respectively. Number of free parameters are shown in the parentheses: **, $p < 4.76E-4$ in paired t-test after Bonferroni correction with 21 pairs of comparison. (B-D) Likelihood of BFQ-learning and RBQ-learning around state-change periods in low (33%-0% and 0%-33%) (B), mid (66%-33% and 33%-66%) (C), and high (100%-66% and 66%-100%) (D) reward-probability states. Trial 0 shows state-change periods. Bold lines and surrounding shaded areas show the means and standard errors of the likelihood. Gray dot straight line shows the significant level of $p = 0.05$ in paired t-test after Bonferroni correction. (E) Summary of likelihood at trial 0, i.e., immediately before state-change period, in Figs. B-D: paired t-test; low, $p = 9.66E-50$; mid, $p = 3.47E-6$; high, $p = 4.54E-76$.

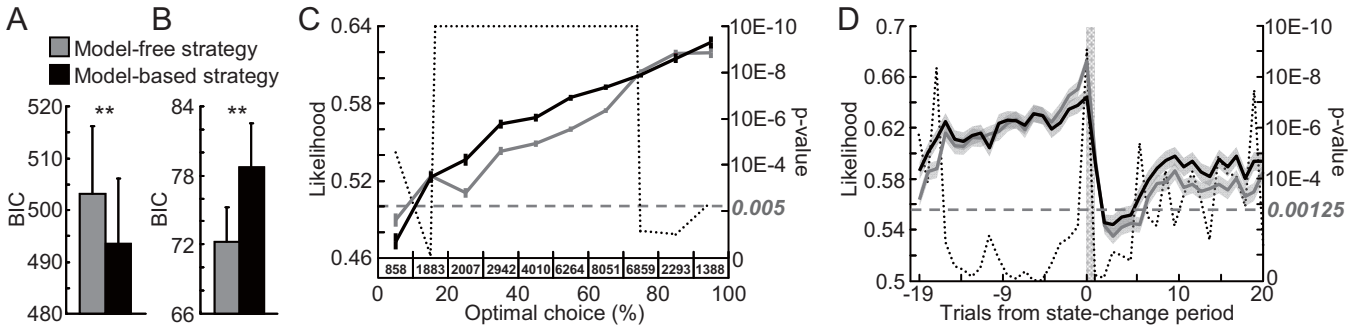

Fig.5. Likelihood of model-free and model-based strategies

(A,B) Model comparison with Bayesian Information Criterion (BIC) in variable-reward (A) and fixed-reward condition (B): **, $p < 0.01$ in paired t-test. (C) Likelihood as a function of optimal choice frequency. Means and standard errors of the likelihood are presented. Dot line shows the significant level of paired t-test which compares the likelihood of model-free and model-based strategies. Table below shows the number of data in each bin. Gray dot straight line shows the significant level of paired t-test. (D) Likelihoods around state-change periods. Data presentations comply with Figs. 4B-D.

task. Here, I conduct a rat's free choice task and analyze the choice behavior with model-free and model-based strategies. I then investigate which strategy best describes the rats' behaviors as a function of rats' behavioral performances.

4.2. Methods

The free choice task consisted of a random sequence of variable-reward- and fixed-reward-condition trials. In each trial, rat selected either left or right hole, and received a reward of food pellet stochastically. A light stimulus was presented only in fixed-reward condition, and thus rats could know the reward condition of each trial. In variable-reward condition, reward probability of each choice changed in 20 to 230 trials among four reward-probability states, i.e., 90%-50%, 50%-90%, 50%-10% and 10%-50%, when a frequency to select an optimal choice reached 80%. By contrast, in fixed-reward condition, reward probabilities were constant in everyday; the probability was 90% and 50% in optimal and non-optimal choice, respectively.

For a model-free strategy, I employed a simple reinforcement learning algorithm, i.e., Q-learning, which gradually updates a value of each choice with past action-reward sequences for deciding choices. In contrast, a model-based strategy assumes that rats already know the four reward-probability states in variable-reward condition, and decides choices based on the state probability. The state probability is updated with hidden Markov model (HMM). I first investigated which strategy better described the choice behaviors in variable-reward and fixed-reward conditions with Bayesian information criterion (BIC). In variable-reward condition, I further investigated the likelihood of both strategies as a function of rats' choice frequencies and trials around state-change periods.

4.3. Results

I used five rats and analyzed 102 sessions of data. Rats' choice behaviors in variable-reward and fixed-reward conditions were significantly fit to model-based and model-free strategies, respectively, as shown in Figs. 5A and B, suggesting that rats employed different learning strategies depending on the context. Further analysis in variable-reward condition showed that rats' behaviors matched model-based strategy when the optimal choice frequency was between 20% and 70%, and when the trials were away from state-change periods, i.e., when rats' choices were fluctuated, as shown in Figs. 5C and D. In contrast, the behaviors matched model-free strategy when the optimal choice frequency was less than 10% and trials immediately before state-change periods. The optimal choice frequency of 10% was usually observed immediately after state-change periods, i.e., over-learning phase. Thus, these results suggest that rats' behavioral strategy shift from model-based to model-free as learning progressed, implying the dynamic system of brain algorithms.

5. Conclusion

Taken together with three experiments, I found that neural activities, parameters and algorithms dynamically changed depending on the sensory stimuli, reward conditions, and learning phases, respectively. My results suggest the existence of sensory-driven and internal-state-driven dynamic system of brain in three different levels, which may be important to adapt highly stochastic environments. Moreover, to test the dynamic brain, machine learning algorithms succeeded to model animals behaviors and to analyze the neural activities. Thus, my study concluded that I developed the machine learning algorithms which could analyze the highly dynamic brain with both top-down and bottom-up ways.

Reference

[1] Funamizu, A., Kanzaki, R. & Takahashi, H. "Distributed representation of tone frequency in highly decodable spatio-temporal activity in the auditory cortex." *Neural, Netw.,* Vol.24, (2011), pp.321-332.

[2] Schultz, W., Preuschoff, P., Camerer, C., Hsu, M., Fiorillo, C. D., Tobler, P. N., & Bossaerts, P. "Explicit neural signals reflecting reward uncertainty." *Phil. Trans. R. Soc. B,* Vol.363, (2008), pp.3801-3811.

[3] Daw, N. D., Niv, Y. & Dayan, P. "Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control." *Nat. Neurosci.,* Vol.8, (2005). pp.1704-1711.