34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Bayesian learning in assisted brain-computer interface tasks

Yin Zhang[1], Andrew B. Schwartz[2], Steve M. Chase[1] and Robert E. Kass[1]

Abstract— **Successful implementation of a brain-computer interface depends critically on the subject's ability to learn how to modulate the neurons controlling the device. However, the subject's learning process is probably the least understood aspect of the control loop. How should training be adjusted to facilitate dexterous control of a prosthetic device? An effective training schedule should manipulate the difficulty of the task to provide enough information to guide improvement without overwhelming the subject. In this paper, we introduce a Bayesian framework for modeling the closed-loop BCI learning process that treats the subject as a bandwidth-limited communication channel. We then develop an adaptive algorithm to find the optimal difficulty-schedule for performance improvement. Simulation results demonstrate that our algorithm yields faster learning rates than several other heuristic training schedules, and provides insight into the factors that might affect the learning process.**

I. INTRODUCTION

Brain-computer interfaces (BCI) promise to restore movement to those who are paralyzed by providing behavioral output directly from the intention to move, bypassing defective neural transmission and muscle activation [3], [5], [7]. However, the control algorithms currently used in BCI are far from perfect: successful decoding still depends critically on the subject's ability to learn how to produce the appropriate neural activity patterns. In closed-loop control, poor estimates of intention can potentially be corrected on-line by the subject. However, to know the appropriate corrective action, the subject has to already have some facility at control. If the subject has no understanding of the mapping between neural activity and effector movement, even the corrective movement will be wrong. To facilitate this learning process, computer assistance is often used at the beginning stages of training to minimize the errors that the subject makes while learning the task [6], [7]. In shared-mode control (SMC), the subject's volitional signal is mixed with a computer-generated "correct" signal to generate the final output. By limiting the errors that the subject produces, SMC increases motivation and keeps the subject engaged in the learning process.

SMC provides a means of directly manipulating the difficulty of the task. How should this difficulty-schedule be chosen? If the difficulty is too low, the errors will be artificially small and the subject will have no pressure to learn. If the difficulty is too high, the errors will be too large to be meaningfully interpreted. The task difficulty must be carefully titrated to the subject's ability to promote rapid learning. Here we introduce a Baysesian framework for modeling a closed-loop BCI learning process that incorporates shared-mode control. By treating the subject as a bandwidth-limited communication channel, we demonstrate an explicit link between the difficulty-schedule and the learning rate. We then develop an adaptive algorithm to find the optimal difficulty schedule for performance improvement. In simulation, our adaptive difficulty-control strategy promotes a marked improvement in learning rate.

*This project was supported by National Institutes of Health Grant RC1NS07031.

[1]Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A.
[2]University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

II. THE TASK

In this paper we consider the center-out reaching task. In this task, starting from the centered home position, the subject needs to control a cursor on a surface to reach a target at position $\boldsymbol{\tau}$. The subject's activity is denoted as φ and the corresponding cursor's movement direction is denoted as ψ. For simplicity, φ is also treated as an angle in this paper. For more complex activity, it can first be projected onto the 1D space. The control system's mapping function f is a rotation with an angle θ^*, i.e.,

$$\psi = f(\varphi;\ \theta^*) = \varphi + \theta^* \qquad (1)$$

where θ^* is the system's parameter and is hidden from the subject. Therefore, the subject's leaning process is essentially a system identification process because in order to get the desired output (the cursor's ideal movement) the subject needs to find θ^* throughout a series of observations about the system's inputs and outputs.

The routine of the subject's knowledge updating at t-th step is shown in Fig. 1 and the notations are shown in Table I. The cursor's desired movement direction ψ_t^* is determined by the cursor's current position $\boldsymbol{s}_t$ and the target's position. The subject's guess about the system parameter, θ_{t-1}, is generated from the subject's current knowledge about θ^*. Together with ψ_t^*, the subject can figure out the desired input φ_t as

$$\varphi_t = f^{-1}\big(\psi_t^*;\ \theta_{t-1}\big) = \psi_t^* - \theta_{t-1} \qquad (2)$$

Correspondingly, the subject's intended output without any assistance is $f(\varphi_t;\ \theta^*)$. However, under SMC, the cursor's actual movement direction is the subject's intended output corrected by a factor λ_t, which is equivalent to the linear combination of the subject's intended output $f(\varphi_t;\ \theta^*)$ and the desired output ψ_t^*, i.e.,

$$\psi_t = \lambda_t f(\varphi_t;\ \theta^*) + (1 - \lambda_t)\psi_t^* \qquad (3)$$

When $\lambda_t = 1$, we have $\psi_t = f(\varphi_t;\ \theta^*)$ which means there is no assistance at all and the subject fully controls the cursor's movement. When $\lambda_t = 0$, we have $\psi_t = \psi_t^*$ which means the

978-1-4577-1787-1/12/$26.00 ©2012 IEEE

TABLE I

NOTATIONS

s_t	the cursor's position at the beginning of the t-th step
ψ_t^*	the cursor's desired movement direction
θ_{t-1}	the subject's guess about the system parameter at the t-th step
φ_t	the system's input, i.e., the subject's activity
ψ_t	the system's output, i.e., the cursor's actual movement direction
$\tilde{\psi}_t$	the subject's perception about ψ_t
λ_t	the weight of the linear combination in SMC

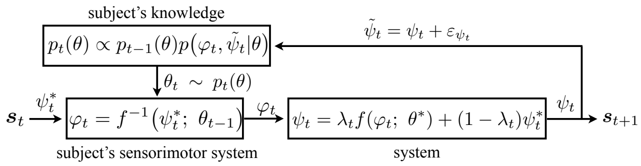

Fig. 1. One Step Updating.

cursor will always move in the ideal direction. Therefore, λ_t reflects the difficulty of the task and our purpose is to adjust λ_t adaptively to facilitate the subject's learning.[1] After the t-th step, the cursor moves to a new position s_{t+1} determined by s_t and ψ_t.

The system's input-output pair $\{\varphi_t, \psi_t\}$ provides some new information about the system and the subject updates the knowledge based on the perception about the input-output pair, denoted as $\{\varphi_t, \tilde{\psi}_t\}$. Here we argue that the subject's perception about the cursor's movement direction, $\tilde{\psi}_t$, is a noisy version of ψ_t, i.e.,

$$\tilde{\psi}_t = \psi_t + \varepsilon_{\psi_t} \tag{4}$$

where ε_{ψ_t} is zero-mean Gaussian noise with variance $\sigma^2_{\psi_t}$. We will discuss this noise term in details in Section III-B.

III. THE MODEL

As we have discussed in the previous section, the learning process in our task is in fact a system identification process and in this paper, we proposed a Bayes learning framework to model this process. In our framework, the subject's knowledge about θ^* is a random variable θ with probability density function $p(\theta)$. At the beginning of the learning task, $p(\theta)$ is fairly flat because the subject is quite uncertain about the system and as the training takes place, the distribution sharps around θ^*.

For simplicity, the subject's initial knowledge $p_0(\theta)$, which is the prior probability before any observations, is assumed as Gaussian distribution, i.e., $p_0(\theta) = \mathcal{N}(\mu_0, \sigma_0^2)$, where μ_0 is an arbitrary guess and σ_0^2 is fairly large.

A. Bayes Learning

The subject's knowledge about θ^* at the beginning of the t-th step, denoted as $p_{t-1}(\theta)$, is the posterior probability after observing the input-output sequence in the first $(t-1)$ steps,

$$p_{t-1}(\theta) = p\big(\theta|\{\varphi_j, \tilde{\psi}_j\}_{j=1}^{t-1}\big). \tag{5}$$

The subject's guess about θ^*, θ_{t-1}, is sampled from $p_{t-1}(\theta)$, i.e., $\theta_{t-1} \sim p_{t-1}(\theta)$. At the end of the t-th step, the subject acquires some new information about θ^* from the perceived data $\{\varphi_t, \tilde{\psi}_t\}$, and updates $p_{t-1}(\theta)$ by Bayes rule,

$$p_t(\theta) \propto p_{t-1}(\theta)p\big(\varphi_t, \tilde{\psi}_t|\theta\big) \tag{6}$$

To get the likelihood, we notice at the t-th step, the system's input is φ_t and the output is ψ_t and θ is the subject's conjecture about the system's parameter. Thus,

$$\psi_t = \varphi_t + \theta \tag{7}$$

and

$$\tilde{\psi}_t = \psi_t + \varepsilon_{\psi_t} = \varphi_t + \theta + \varepsilon_{\psi_t} \tag{8}$$

Since $p\big(\tilde{\psi}_t, \varphi_t|\theta\big) = p\big(\tilde{\psi}_t|\varphi_t, \theta\big)p\big(\varphi_t|\theta\big)$ and φ_t is independent of θ, we have the likelihood as

$$p\big(\tilde{\psi}_t, \varphi_t|\theta\big) \propto \mathcal{N}\big(\varphi_t + \theta,\ \sigma^2_{\psi_t}\big) \tag{9}$$

B. Noise Term ε_{ψ_t}

To get ε_{ψ_t}, in this paper we treat the subject as a communication channel with a limited bandwidth. Form Shannon-Hartley theorem we know that the channel capacity C_{max}, or, in our case, the upper bound on the information rate the subject can acquire each step, is

$$C_{max} \geq C_t = B\log\left(1 + S_t/N_t\right) \tag{10}$$

where C_t is the information rate at the t-th step, B is a constant related to the bandwidth of the channel, S_t and N_t are the powers of the signal and the noise at the t-th step.

To get S_t, we notice when the system's input is φ_t, the subject's expected cursor's movement direction is ψ_t^*. Thus, the difference between the subject's expected output and the actual output, i.e., the error, can be considered as the signal power, i.e.,

$$S_t = (\psi_t - \psi_t^*)^2 \tag{11}$$

The noise power comes from the noise term ε_{ψ_t}. Since it is Gaussian noise, the noise power equals to its variance

$$N_t = \sigma^2_{\psi_t} \tag{12}$$

Thus, we get the information rate at the t-th step as

$$C_t = B\log\big(1 + (\psi_t - \psi_t^*)^2/\sigma^2_{\psi_t}\big) \tag{13}$$

In this paper, we assume the information rate can reach its upper bound, i.e., $C_t = C_{max}$, then, we have the expression of the noise's variance as

$$\sigma^2_{\psi_t} = \alpha(\psi_t - \psi_t^*)^2 \tag{14}$$

where $\alpha = \exp\left(C_{max}/B\right) - 1$ is a constant determined by the subject's capability. Large α indicates the high capability of acquiring new information and vice versa. From Eq. 14, we can see the variance of the subject's perception is proportional to the error. What's more, notice the error can be expressed as

$$(\psi_t - \psi_t^*)^2 = (\theta^* - \theta_{t-1})^2\lambda_t^2 \tag{15}$$

Thus after the subject's guess θ_{t-1} is generated, the error is solely determined by the task's difficulty λ_t. This result reveals that when SMC is minimal ($\lambda_t = 1$), the subject

[1] Actually λ_t is the percentage of task difficulty, not the absolute amount. The absolute task difficulty is determined by the experimental design.

receives veridical feedback about its movement, with the error that provide the most useful information for updating its internal conception of the motor transform. However, large error also leads to increased perceptual noise, ε_{ψ_t}, limiting the subject's ability to utilize the feedback. This conclusion is similar to the statement in [2], where authors argued that there is a maximum volume of new information the subject can acquire each step. If the information provided by the task exceeds this threshold, more information will harm the subject's perception. [1] also demonstrates the dependence of the degree of specificity on the task difficulty. This tension suggests that optimal learning may be driven by intermediate levels of assistance that decrease as the subject gains proficiency at the task. Finding the proper schedule of λ_t to optimize the rate of learning is the goal of this work.

C. Posterior Probability Updating

With the expression of likelihood, we can update the posterior probability, which is the subject's knowledge about θ^* after the t-th step. Since p_0 is Gaussian and the likelihood at each step is also Gaussian, from induction we know $p_t(\theta)$ is Gaussian. Assuming $p_{t-1}(\theta) = \mathcal{N}\left(\mu_{t-1}, \sigma_{t-1}^2\right)$, we have the updated posterior as $p_t(\theta) = \mathcal{N}\left(\mu_t, \sigma_t^2\right)$, where

$$\mu_t = \left(\mu_{t-1}\sigma_{t-1}^{-2} + \left(\tilde{\psi}_t - \varphi_t\right)\sigma_{\psi_t}^{-2}\right)\sigma_t^2 \quad (16)$$

$$\sigma_t^2 = \left(\sigma_{t-1}^{-2} + \sigma_{\psi_t}^{-2}\right)^{-1} \quad (17)$$

Some Observations: From the above equations, we can see the updated posterior mean is the linear combination of μ_{t-1}, which comes from the previous knowledge, and $(\tilde{\psi}_t - \varphi_t)$, which comes from the new precieved data. From Section II, we can expand $(\tilde{\psi}_t - \varphi_t)$ into two parts as

$$\tilde{\psi}_t - \varphi_t = \left(\lambda_t\theta^* + (1-\lambda_t)\theta_{t-1}\right) + \varepsilon_{\psi_t} \quad (18)$$

The first part $\left(\lambda_t\theta^* + (1-\lambda_t)\theta_{t-1}\right)$ reflects how much information about θ^* the current step provides. When λ_t is close to 1, this part is close to θ^* which means much information about the ground-truth parameter θ^* is provided. On the other side, when λ_t is close to 0, this part is close to θ_{t-1}, which means the subject's observation is quite similar to what it has already learnt and little information about θ^* is provided. So, from this point of view, to make the subject learn as much as possible, λ_t should be set as large as possible.

The second part ε_{ψ_t} is a random variable with variance $\sigma_{\psi_t}^2$. From the discussion in Section III-B we know when λ_t is large, with high probability ε_{ψ_t} is far away from 0. So, from this point of view, we want to keep λ_t small. Therefore, we hope to find a balance between those two parts.

IV. Adaptive Difficulty Control

In this section, we design a strategy which can automatically adjust the task's difficulty at each step so that the subject can learn as fast as possible. Specifically, when the system obtains the subject's input φ_t, we hope the cursor can move in a proper direction that helps the subject improve its knowledge as much as possible.

To do this, we first define the risk of the subject's current knowledge as the mean squared error between θ and θ^*, i.e.,

$$R(p_t) = E_{p_t(\theta;\mu_t,\sigma_t)}(\theta - \theta^*)^2 = (\mu_t - \theta^*)^2 + \sigma_t^2 \quad (19)$$

where $(\mu_t - \theta^*)^2$ is the bias, which measures the distance between the subject's knowledge and the ground-truth, and σ_t^2 is the variance, which measures the confidence about subject's knowledge.

To make the risk converge to 0 as fast as possible, one heuristic way is to minimize the expected risk of the next step, i.e.,

$$\lambda_t = \arg\min_\lambda E_{\varepsilon_{\psi_t}}\left[R(p_t)\right] \quad (20)$$

where p_t is given by Eq. 16.

However, from the simulation results (Fig. 4) we find this intuitive method doesn't work very well. It is because at the first few steps, the optimal λ minimizing Eq. 20 focuses on decreasing the variance of the risk while keeping the bias high. In this case, the subject will be quite confident about some wrong knowledge after a few initial steps. Thus, more steps are needed to correct it. To prevent this case, instead of minimizing the expected risk, we try to minimize the expected bias while keeping the variance untouched. Replacing $R(p_t)$ in Eq. 20 by the bias term $(\mu_t - \theta^*)^2$, we have the new optimization problem,

$$\lambda_t = \arg\min_{\lambda\in[0,1]} E_{\varepsilon_{\psi_t(\lambda)}}\left[\mu_t(\lambda) - \theta^*\right]^2. \quad (21)$$

V. Simulation

In our simulation, all angles are confined in $(-\pi, \pi]$. Thus, the distribution of the subject's knowledge $p_t(\theta)$ is the wrapped Gaussian distribution, $p_t(\theta) = \sum_{j\in\mathbb{Z}} q_t(\theta + 2\pi j)$, where $q_t(x) = \mathcal{N}(\mu_t, \sigma_t^2)$. And the distance between two angles θ and θ^* is defined as $\min_{j\in\mathbb{Z}}(\theta - \theta^* + 2\pi j)^2$.

At the beginning, the subject's knowledge p_0 is an uniform distribution on $(-\pi, \pi]$. It is equivalent to the wrapped Gaussian distribution with infinity variance and in the first updating the subject will fully trust the likelihood.

The home position is $(0, 0)$ and the target is randomly posed on a circle with radius 50. The cursor's step length is 1. We repeat 100 times of the learning process and each time the system parameter θ^* is uniformly sampled from $(-\pi, \pi]$. The results shown are averaged over those 100 trials.

The learning curves of risk defined in Eq. 19 corresponding to different difficulty control strategies are compared. We also compare the subject's actual error, which is the difference between the subject's intended output and the desired direction, i.e., $L_t = \left(f(\varphi_t;\ \theta^*) - \psi_t^*\right)^2$. Since the results of L_t are similar to those of risk, we do not show the comparison here.

A. Results

In the first simulation experiment, we fix the task difficulty λ throughout the whole process to study the properties of the proposed learning framework. The results are shown in Fig. 2. Here we consider three values of α, small one ($\alpha = e^0$), median one ($\alpha = e^2$), large one ($\alpha = e^4$) and three values of λ, easy one ($\lambda = 0.2$), median one

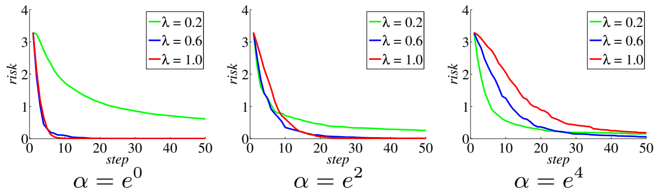

Fig. 2. Comparison between learning curves under fixed difficulty scheduling.

$(\lambda = 0.6)$, hard one $(\lambda = 1)$. As we have seen in Sec. III-B, when α increases, the capability of the subject acquiring new information decreases. Thus $\alpha = e^0$ corresponds to the case where the subject can acquire the most information. In this case, simply setting $\lambda = 1$ will make the subject learn fastest, just as shown in Fig. 2. As α increases, the subject's capability of acquiring new information decreases and $\lambda = 1$ becomes too difficult for the subject. Thus, at the beginning of the learning process, the convergence rate is quite slow. Therefore, if the subject's capability is not very strong, making the task easy when the subject knows little about the system is more helpful. This result agrees with the intuition and the statement of [1], [2].

To demonstrate the effectiveness of our proposed difficulty control strategy (denoted as ADP-B), we first compare it with the strategy of fixed difficulty (denoted as FIX). The results are shown in Fig. 3. The blue curves correspond to the learning curves under fixed difficulty and the red one corresponds to our adaptive strategy. The averaged λ trajectory under our strategy is also shown in figures as the red dash curves. From the results we can see that our strategy is almost always better than any fixed difficulty scheduling. That demonstrates the adaptive difficulty control is necessary for fast learning and our strategy provides a good choice.

Finally, we compare ADP-B with other two control strategies. The first one (denoted as AVGTRJ) is using the averaged λ trajectory obtained from ADP-B universally. The second one (denoted as ADP-R) is choosing λ to minimize the subject's expected risk at each time point as discussed in IV. The results are shown in Fig. 4. The blue dash curve corresponds to the averaged λ trajectory under ADP-R and the red dash curve corresponds to the averaged λ trajectory under ADP-B and AVGTRJ. The deficiency of AVGTRJ compared to ADP-B demonstrates there exists no universal difficulty scheduling that can work well on all trials and the optimal control strategy should be adaptive on different trials.

Just as the discussion in Section IV, ADP-R performs much worse than ADP-B from the results. We find that ADP-R initially decreases quickly, but converges relatively slowly. From Section IV we know this is because ADP-R focuses on reducing the variance of the first several steps. ADP-B, as a greedy heuristic to avoid this problem, results in fairly fast reductions in overall risk, while still allowing rapid convergence of risk to optimal levels.

VI. CONCLUSIONS

The Bayesian learning framework developed here has two novel features that capture the specifics of learning in a

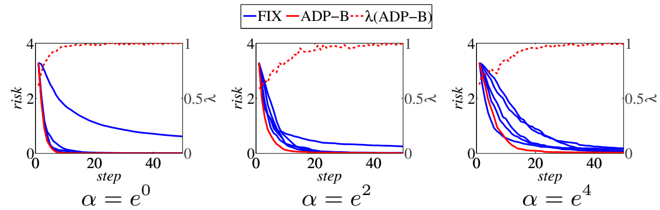

Fig. 3. Comparison between learning curves under adaptive difficulty scheduling and fixed difficulty scheduling. λ trajectories are shown as dash curves.

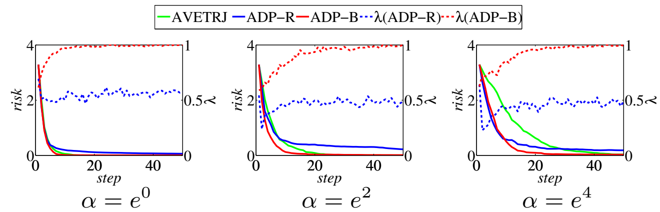

Fig. 4. Comparison between learning curves under adaptive difficulty scheduling and universal difficulty scheduling. λ trajectories are shown as dash curves.

BCI context. First, we explicitly incorporate the shared-mode control process used by Schwartz and colleagues to assist in subject training [6], [7]. Second, we treat the subject as a band-limited communication process. The resulting training schedule adjusts task difficulty to improve the subject's learning rate. Our simulation results demonstrate the effectiveness of the adaptive training strategy.

We developed our framework by analogy of the subject as a limited-bandwidth communication channel, which links the perceptual noise in the system to the overall success rate: the greater the error in the subject's output, the larger the noise in the observed movement. An identical framework results if one instead assumes that the subject's motivation depends on overall success rate. Because motivation affects attention, and attention impacts perceptual noise [4], increased computer assistance leads to decreased perceptual noise. From this perspective our simulations show that learning rate improves when task difficulty is manipulated adaptively while explicitly accounting for subject motivation. These promising results suggest that Bayesian learning can offer useful insights and methods for teaching subjects to use a brain-computer interface device.

REFERENCES

[1] M. Ahissar and S. Hochstein. Task difficulty and the specificity of perceptual learning. *Nature*, 387(22):401–406, 1997.
[2] M. A. Guadagnoli and T. D. Lee. Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36(2):212–224, 2004.
[3] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442:164–171, 2006.
[4] Z. L. Lu and B. A. Dosher. External noise distinguishes attention mechanisms. *Vision Research*, 38:1183–1198, 1998.
[5] A. B. Schwartz, X. T. Cui, D. J. Weber, and D. W. Moran. Brain-controlled interfaces: movement restoration with neural prosthetics. *Neuron*, 52:205–220, 2006.
[6] D. M. Taylor, S. I. Tillery, and A. B. Schwartz. Direct cortical control of 3D neuroprosthetic devices. *Science*, 296:1829–1832, 2002.
[7] M. Velliste, S. Perel, M. C. Spalding, A. S. Whitford, and A. B. Schwartz. Cortical control of a prosthetic arm for self-feeding. *Nature*, 453:1098–1101, 2008.