 DOI:10.3217/978-3-85125-260-6-125

Probabilistic Co-Adaptive Brain-Computer Interfacing

M. J. Bryan[1], S. A. Martin[2], W. Cheung[3], R. P. N. Rao[1,2]

[1]Neural Systems Laboratory, Department of Computer Science, University of Washington, Seattle, WA, USA; [2]Center for Sensorimotor and Neural Engineering, University of Washington, Seattle, USA; [3]Graduate Training Center of Neuroscience, International Max Planck Research School, University of Tübingen, Tübingen, Germany

Correspondence: Rajesh P. N. Rao. E-mail: rao@cs.washington.edu

Abstract. Brain-computer interfaces (BCIs) are confronted with two fundamental challenges: (a) the uncertainty associated with decoding noisy brain signals, and (b) the need for co-adaptation between the brain and the interface to account for non-stationarities that emerge over time. We introduce a new approach to brain-computer interfacing that addresses these challenges using the framework of partially observable Markov decision processes (POMDPs). POMDPs provide a principled way to handle uncertainty and achieve co-adaptation: (1) Bayesian inference is used to compute posterior probability distributions ("beliefs") over brain and environmental state, and (2) actions are selected based on belief distributions so as to maximize total expected reward. By employing methods from reinforcement learning, the POMDP's reward function can be updated over time to allow for co-adaptive behavior. We illustrate our approach using a simple non-invasive BCI system which optimizes the speed-accuracy trade-off for subjects based on the amount of uncertainty measured in their brain signals. Our results demonstrate that the POMDP-based BCI can detect changes in user brain state and co-adaptively switch control strategies on-the-fly to maximize reward.

Keywords: Probabilistic BCI, Co-Adaptive Control, SSVEP, EEG, POMDP

1. Introduction

We propose a new approach to BCI design that combines Bayesian inference over brain states with an optimzing control scheme for selecting actions. The resulting framework, which is based on POMDPs [Cassandra et al., 1994], allows the BCI to handle uncertainty in decoding as well as to co-adapt with the user to converge to optimal task strategies. The proposed approach generalizes to a wide variety of applications, including both invasive and non-invasive BCIs for communication and robot control.

2. Methods

The POMDP framework allows us to model a decision process in which an agent acts in the world to affect outcomes, which are rewarded based on a reward function. In this model, the agent views the world indirectly through an observation function. We illustrate the use of this model for deciding when to collect more information from the brain and when to stop and make a decision about the user's intent, similar to [Park et al., 2011]. The relative weights of the users' preference for accuracy versus expediency are expressed through the reward function. We demonstrate the framework using an EEG BCI based on steady state visually evoked potentials (SSVEPs).

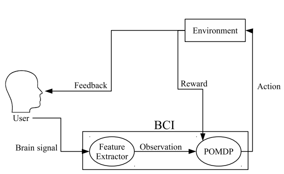

Figure 1: System flowchart

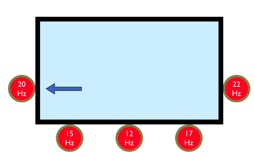

Figure 2: *During data collection phases, the user sees cues indicating which stimulus to fixate on*

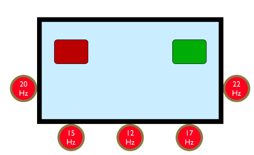

Figure 3: *During the co-adaptive experiment, the user's goal is to select the 'green box' stimuli*

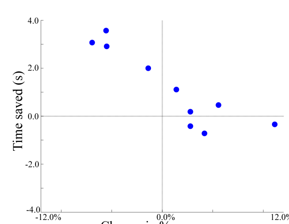

Figure 4: *Plot comparing average POMDP's performance to fixed-window classification for each user*

The reward function implicitly encodes the mapping between the various brain states we are tracking and control. For instance, in a task involving only movement left and right, and an SSVEP-type BCI with 5 channels, the reward function would tell our agent which action to take (moving left or right) given a particular brain state it may be in (the

 DOI:10.3217/978-3-85125-260-6-125

choice of SSVEP channel). Through reinforcement learning, we allow this control mapping to be non-stationary. In our experiment, reinforcements are given implicitly since the task has a pre-defined failure mode. We note that this signal could also be derived in a task-independent way using user perception of system failure [Gürel and Mehring, 2012]. This reinforcement process allows the user the freedom to *search the space of possible brain states* for those that work best for them. The result is a BCI which adapts to the user's behavior, who is also adapting to the BCI and the present situation (see Fig. 1).

We train the POMDP's observation model using labeled data. We collect the data through a training paradigm where the subject focuses on each flashing LED in a random order (Fig. 2). The POMDP's policy is then approximated using apre-packaged method (see [Kurniawati et al., 2008]). We present results from two experiments. In the first, the user provides us a data set using the same paradigm as our training data collection. In post-processing, we fix the POMDP's reward function and then run it in simulation to examine its trade-off between speed and accuracy.

In the second experiment, we present the user with a pair of targets. Their goal is to hit the green target. The user is free to pick any mapping between one of the SSVEP frequencies and a target. The BCI automatically modifies the reward function depending on the success of the users control in each trial (see Fig. 3).

3. Results

Overall, we saw the POMDP model gave users either an increase in accuracy, a decrease in decision time, or both, for one fixed parameterization. We should note that in practice the reward function would be tuned to the particular user's needs. We found that varying the reward function (e. g., changing the penalty for waiting an additional time step) allows one to trade off time spent collecting data with the expected accuracy of the decision.

An additional advantage of this model is that the BCI adjusts its behavior according to the expected usefulness of future observations. If the observation function indicates a particularly noisy brain signal, the BCI automatically accepts making decisions with less confidence.

In experiment 2, we found the co-adaptive system adjusted to non-stationarities imposed by the user's changing control strategies. Consider Fig. 5 which depicts the effect of changing control mappings for one of the users. Because some channels were difficult to classify for this user, they decided to change their mapping to use the 12 and 15 Hz stimuli. At time τ in Fig. 5, the user swapped their control mapping but the POMDP BCI was able to recognize this change in the user's strategy through reinforcement and it co-adapted with the user over several trials to converge to an appropriate policy.

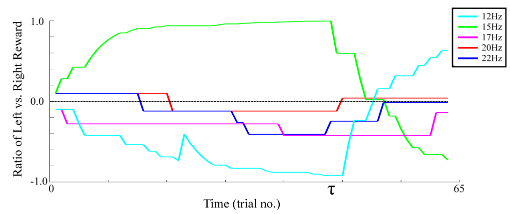

Figure 5: *Co-adaptive convergence of the POMDP BCI for User 1's 5-channel experiment. At first, the user tried various channels but found that 12 Hz and 15 Hz had the most accurate performance. Occasional misclassifications resulted in other channels changing as well. At trial τ, the user switches the mapping of 12Hz and 15Hz.*

4. Conclusion

Our results suggest that the POMDP framework is well-suited to handling uncertainty in BCI systems and provides a principled approach to achieving co-adaptation between the brain and the BCI. Our results from an SSVEP BCI demonstrate that this model offers the ability to choose the control scheme the user finds most intuitive in a given situation. It also provides a balance between accuracy and speed in a principled manner.

5. Acknowledgments

This research was supported by ARO Award no. W911NF-11-1-0307, NSF award no. 0930908, the Office of Naval Research (ONR), the Center for Sensorimotor and Neural Engineering, and a Mary Gates Research Scholarship.

References

Cassandra, A., Kaelbling, L., and Littman, M. (1994). Acting optimally in partially observable stochastic domains. In *Proceedings of Twelfth National Conference on Artificial Intelligence (AAAI)*, Seattle, Washington, USA.

Gürel, T. and Mehring, C. (2012). Unsupervised adaptation of brain machine interface decoders. *CoRR*, abs/1206.3666.

Kurniawati, H., Hsu, D., and Lee, W. (2008). SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces. In *Proceedings of Robotics: Science and Systems IV*, Zurich, Switzerland.

Park, J., Kim, K., and Song, Y. (2011). A POMDP-based optimal control of P300-based brain-computer interfaces. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI) NECTAR Track*.