Automata Models in BCI's

Omar AbdAlAal
dept. of Computer Engineering
Ain Shams University
Cairo, Egypt
18p2998@eng.asu.edu.eg

Ahmed Mohamed ElMahdy
dept. of Computer Engineering
Ain Shams University
Cairo, Egypt
2100723@eng.asu.edu.eg

Hassan Hassan
dept. of Computer Engineering
Ain Shams University
Cairo, Egypt
18p1726@eng.asu.edu.eg

Abdelrahman Ghazy
dept. of Computer Engineering
Ain Shams University
Cairo, Egypt
19p6651@eng.asu.edu.eg

Abanoub Henen
dept. of Computer Engineering
Ain Shams University
Cairo, Egypt
21p0356@eng.asu.edu.eg

Omar Mahmoud
dept. of Computer Engineering
Ain Shams University
Cairo, Egypt
21p0208@eng.asu.edu.eg

***Abstract*—The utilization of Brain-Computer Interfaces (BCIs) has revolutionized the interaction between human cognition and technological systems, especially for individuals with severe disabilities. This research addresses the challenge of processing diverse neural signals obtained through invasive, semi-invasive, and non-invasive methods, including EEG and fMRI. Hidden Markov Models (HMMs) emerge as a potent probabilistic tool for modeling the sequential and dynamic characteristics of brain signals. Drawing inspiration from automata theory, HMMs encapsulate hidden states inferred from observable outputs and employ probabilistic transitions for decoding mental states. The research explores key signal processing techniques such as Independent Component Analysis (ICA), feature extraction methods, and classification algorithms, alongside optimization methods like genetic algorithms, and compares and contrasts them against HMMs. Despite their potential, challenges limit their widespread application. Advancing neurophysiological understanding and addressing these challenges hold promise for enhancing the efficacy of HMM-based BCIs.**

***Index Terms*—Brain-Computer Interface, Hidden Markov Model, Automata, Automata Applications**

I. Introduction

The primary challenge lies in processing brain signals collected through different neural interfaces: invasive methods like intracortical microelectrodes (IM), semi-invasive methods like electrocorticography (ECoG), and non-invasive methods like electroencephalography (EEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI) [3].

To tackle this, a range of signal-processing techniques is utilized, including feature extraction methods like independent component analysis (ICA), wavelet transformations, and autoregressive modeling. Classification algorithms, such as support vector machines (SVM), alongside machine learning models like artificial neural networks, are also essential tools. In addition, optimization techniques, such as genetic algorithms or particle swarm optimization, and pattern recognition approaches—like hidden Markov models (HMMs)—offer powerful solutions.

This research focuses on HMMs, which can be thought of as a specialized form of automata. Similar to classical automata, HMMs have states and transitions. However, the states in HMMs are "hidden," inferred through observable outputs, while the transitions themselves are governed by probabilities. HMM's can be represented as classical FSM's, with probabilities in the transition function. The output is nondeterministic or overlapping in nature.

There are several reasons why HMMs are suitable for processing brain signals in BCI applications.[2]

- **Sequential Data Modeling:** Brain signals, such as those from EEG or MEG, have a temporal structure, meaning they change over time. HMMs are designed to handle sequential data by modeling transitions between states, which makes them ideal for capturing the dynamic nature of neural activities.
- **Hidden States Representation:** Brain signals are influenced by latent mental or cognitive states that cannot be directly observed. HMMs infer these hidden states from observable signal features, effectively bridging the gap between raw data and meaningful mental representations.
- **Probabilistic Nature:** Neural signals are inherently noisy and complex, influenced by various factors like overlapping events. HMMs account for uncertainty by using probabilities for state transitions and observations, making them robust against variability in signal data.
- **Temporal Dependencies:** Unlike other models that treat data points independently, HMMs consider the dependency of a current state on the previous

one. This Markovian property aligns well with how brain processes evolve over time, enabling accurate modeling of tasks like motor imagery or mental state transitions.

II. Analysis

As a probabilistic model designed to handle sequential data, an HMM models a brain signal as a series of hidden states—with each state representing, for example, a distinct mental or cognitive state—and learns the transition probabilities between these states. This makes HMMs particularly suited for tasks where the temporal evolution of the signal is important, such as decoding the sequence of neural patterns corresponding to a user's intent.

The model is then able to predict the user's intention from the observed data and use it to control a computer system or robotic device. In addition, HMMs can be used for decoding motor imagery activities, which allow users with severe disabilities to interact with computers using only their minds.

The HMM method, based on the Bayesian posterior probability maximization approach, has been successfully employed to classify time series, with states at any given time t being affected by states at the previous time (t - 1). This technique has shown its utility in resolving speech recognition problems and is widely used for signal analysis, classification, modeling, and control purposes.

The solution of three fundamental problems is required to construct a Hidden Markov Model (HMM) for the given sequence of observed states:

1) The evaluation problem can be stated as follows: Given an HMM with transition probabilities a_{ij} and b_{jk}, determine the probability that a particular sequence of visible states (VT) was generated by this model.
2) Decoding problem. Given an HMM and a set of observations (V^T), we need to determine the most probable sequence of hidden states ω^T that result in these observations.
3) The learning problem. Given an enlarged structure of the model with a specified number of states and visible states but without knowledge of transition probabilities a_{ij} and b_{jk}, learning can be performed by determining the most plausible model from a training sample of visible states.

HMMs can be utilized in BCIs as probabilistic automata that calculate the likelihood of a given sequence of feature vectors. Each state of the automaton models the probability distribution for observing a particular feature vector, with Gaussian models being commonly used in BCI applications.

The use of HMMs for classifying time series has been demonstrated to be effective due to their inherent nature. As the EEG contains distinct features that can be distinguished in the temporal domain, the HMM method was utilized for EEG classification in BCI. However, HMMs have not seen widespread application in BCI development yet, despite the relatively high success rates reported by known studies. One major obstacle is the necessity of identifying an invariable set of observable states related to an event, which may prove difficult when analyzing EEG signals; for example, while studying induced desynchronization across different channels. If signals associated with close localized events and similar dynamics are being analyzed, it becomes hard to identify stable observable states; requiring instead the identification of process attractors related only to classified events without including background nervous system activity. A further exploration into EEG signals and neurophysiological brain functioning fundamentals could possibly provide such an opportunity, thus increasing the relevance of using HMMs significantly.

Challenges of Using HMMs in BCIs

A. EEG Signal and Data Challenges

- **Noise:** EEG recordings are highly susceptible to artifacts (eye movements, muscle, and environmental noise) and have a low SNR. These factors can violate HMM assumptions and degrade performance. [2]
- **Non-stationarity:** Brain signals change over time and across sessions. HMM-based analyses often assume that observations within each "state" have stationary statistics, but EEG is inherently nonstationary. In practice, data must be segmented into short epochs that are assumed stationary for the HMM to work. As one recent model paper notes, "physiological data are nonstationary," so EEG must be windowed for HMM training. This segmentation adds complexity and may degrade performance if brain dynamics evolve continuously.
- **High-dimensionality:** EEG data can have many channels and frequency bands, leading to large feature vectors. HMMs with Gaussian mixture emissions must estimate many parameters when using high-dimensional features. To avoid computational intractability, low-dimensional features or dimensionality reduction are typically needed.

B. Model & Theoretical Challenges

- **Choice & Interpretability of Hidden states:** An HMM requires selecting the number of hidden states and their structure, but physiological brain states are not directly observed. Using a fixed number of states may be too simplistic. One recent study pointed out that forcing a predetermined state count "may not capture the heterogeneous neural dynamics across individuals". In other words, fixed-state HMMs can mischaracterize individual differences in EEG patterns. Moreover, hidden states are often hard to interpret in terms of cognitive or neural phenomena.

C. Practical & Real-Time Challenges

- **Calibration & Adaptation:** Most HMM-based BCI systems require per-user calibration: training the model on each subject's data. Inter-subject variability in EEG means an HMM trained on one person may not generalize to another. Likewise, even within a session, EEG patterns drift, necessitating frequent retraining or adaptation. Addressing this is challenging. One survey notes that achieving "calibration-free" BCIs is an open problem and suggests domain adaptation or transfer learning as possible remedies.

III. Related Works

There are six widely used brain signal processing techniques in BCIs. *Independent Component Analysis (ICA), Wavelet Transformations & Autoregressive Modeling, Support Vector Machines (SVM), Hidden Markov Models (HMM), Neural Network Algorithms,* and *Genetic Algorithms (GA) & Particle Swarm Optimizations (PSO).*

Each technique is used in a separate stage of brain signal processing. And some are overlapping, and can be used in place of one another depending on the goal & application.

1) **ICA:** is a source separation technique, used in the signal pre-processing stage. When recording brain signals (like EEG), the data obtained is a complex mixture of various neuronal sources as well as noice from artificats (like eye blinks, muscle movements, or environmental interference). ICA decomposes the multichannel signals into statistically independent components. This separation helps isolate the true brain activity from these unwanted artifacts and improves the signal-to-noise ratio, setting the stage for more accurate downstream analysis. Essentially, ICA "cleans up" the data so that further processing, whether classification or decoding, can focus on the relevant neural signals.

2) **Wavelet Transformations & AR Modeling**
   - *Wavelet Transformation:* decompose a non-stationary signal like EEG into its time-frequency components. Unlike Fourier transforms, they capture transient features and local variations in both time and frequency domains. This multiresolution capability makes wavelets ideal for identifying brief changes or bursts in brain signals that may indicate cognitive or motor processes. In many practical applications, wavelet packet decomposition is used to segment the signal into sub-bands, providing a detailed frequency analysis.
   - *AR Modeling:* Once the signal has been decomposed into various frequency bands, AR modeling comes into play as a method for capturing the dynamic characteristics of each component. Essentially, an AR model predicts future signal values based on past observations. When applied to each wavelet-derived sub-band, the resulting AR coefficients serve as robust features that summarize the temporal dynamics of that frequency component. These features often feed into classifiers, improving the accuracy of task recognition or mental state distinctions in BCIs.

   Together, wavelet transform and AR modeling form a powerful feature extraction pipeline. The wavelet transform handles the non-stationary and multi-scale nature of EEG signals, while AR modeling parameterizes the signal's temporal behavior. This combination produces a detailed, high-resolution representation of the EEG data before any classification or decoding algorithm is applied.

3) **SVM:** are used as discriminative classifiers. Once relevant features are extracted from brain signals, the SVM takes a fixed feature vector and finds the optimal separating hyperplane between classes. SVMs offer good classification performance even with limited training data, are relatively straightforward to tune, and have proven effective across numerous BCI applications such as motor imagery classification, where each trial is treated as a separate sample.

4) **HMM:** are generative probabilistic models designed to capture temporal dynamics. They assume that the observed brain signals are produced by a sequence of hidden states; the model learns both the probability of transitions between these states and the likelihood of observing a given feature set from each state. This sequential modeling is particularly advantageous for tasks that involve dynamic changes over time, such as tracking mental state transitions during an ongoing BCI task.

   Rather than treating each feature vector as an

isolated instance, HMMs naturally incorporate the ordering of data. This allows them to model temporal dependencies, making them better suited for continuous decoding of brain signals where the evolution of states (like initiation, maintenance, and termination of motor imagery) is critical.
HMMs provide interpretable probabilistic state sequences and can effectively model the evolution of "hidden" neural states over time. This makes them ideal for real-time tracking and decoding in BCIs, addressing tasks where the temporal context directly impacts performance.

5) **Neural Networks:** focus on learning complex, non-linear mappings directly from data. They "learn" representations that best capture the relationships within input features (like processed EEG signals) through data-driven training using backpropagation.
Many NN-based approaches are designed to both extract features and perform classification in a single, end-to-end system. For example, Convolutional Neural Networks (CNNs) can capture spatial patterns in EEG topographies, while recurrent (RNNs) or long short-term memory (LSTM) networks address temporal dependencies when processing sequences of data. This integrated feature learning often allows neural networks to outperform traditional methods when large, well-labelled datasets are available.
One challenge with neural networks is their "black-box" nature. While recent advances in explainable AI are helping bridge this gap, NN models often provide less direct interpretability regarding the sequential or state-based nature of brain processes.

6) **GOs & PSOs:** re typically used for tuning parameters or selecting features to optimize parts of the BCI pipeline. For example, they might optimize the filter parameters, spatial channel selections, or even the hyperparameters of a classifier (which could be an HMM or another model). Their strength lies in navigating high-dimensional, non-linear optimization problems where traditional methods might get stuck in local minima.

- *GAs:* Mimic the process of natural selection by using operations like crossover, mutation, and selection to evolve a population of candidate solutions toward better performance.
- *PSO:* Inspired by the social behavior of swarms, PSO involves a group of candidate solutions (particles) that adjust their positions in the search space based on both their own experience and the success of their peers.

TABLE I
COMPARISON OF TECHNIQUES FOR PROCESSING BRAIN SIGNALS IN BCIS

Technique	**Key Role**	**Typical Usage**	**Application**	**Difference from HMM**
ICA	Isolates neural components by performing blind source separation.	Primarily used in pre-processing to remove artifacts (e.g., ocular, muscular) from EEG signals.	Enhances signal quality in cognitive studies by separating noise from true brain activity.	A pre-processing, stateless method, unlike HMM, which models temporal state transitions.
Wavelet Transformations & AR Modeling	Extracts time-frequency features and models signal dynamics.	Used for decomposing non-stationary brain signals and capturing both spectral and temporal information.	In epilepsy research, transient spectral changes detected via wavelets (coupled with AR modeling for dynamics) assist in seizure prediction.	Provides feature extraction without explicitly modeling sequential Markovian state transitions.
SVM	Acts as a classifier by constructing optimal separating hyperplanes.	Often employed after feature extraction to classify brain states from static, high-dimensional features.	Used in P300 speller systems to distinguish between intended and non-intended cognitive responses.	A static classifier, whereas HMM is generative and captures temporal dynamics.
HMM	Models temporal dynamics by representing the sequence of hidden cognitive or motor states.	Applies probabilistic techniques (e.g., Baum–Welch for training, Viterbi for decoding) to infer sequential state transitions from EEG signals.	Widely used in controlling prosthetic limbs by decoding continuous EEG signals to track user intent over time.	Focus of this research; excels in capturing temporal dependencies.
Neural Network Algorithms	Provides non-linear mapping and learns complex patterns directly from raw data.	Utilized for end-to-end classification and regression tasks by learning hierarchical representations of brain signals.	In motor imagery BCI systems, CNNs have been used to automatically extract features for intent detection.	Does not assume a Markov process and requires more data to learn temporal structures.
GA & PSO	Optimize parameters and perform feature selection in high-dimensional spaces.	Deployed as metaheuristic optimizers to tune hyper-parameters (even within HMMs) or select optimal feature sets.	In online BCI implementations, these methods help fine-tune spatial filters and model settings to improve response times.	Serve as optimization tools rather than direct modeling techniques.

IV. Conclusions & Future Work

In this paper, we have examined the strength of HMMs in processing brain signals for BCIs. Our analysis demonstrates that HMMs are particularly strong in capturing the temporal and stochastic nature of neural data by modeling the underlying latent states and their probabilistic transitions. By incorporating sequential dependencies and effectively handling the inherent noise in neural recordings, HMMs can decode complex sequences such as motor imagery tasks and other cognitive activities. Alongside techniques like Independent Component Analysis and Support Vector Machines, HMMs provide a comprehensive framework for real-time applications, even though challenges such as establishing stable observable states persist.

Building on these findings, several areas for future research emerge:

1) **Hybrid Modeling Approaches:** Integrating HMMs with deep learning techniques—such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs)—could leverage the strengths of both probabilistic reasoning and data-driven feature learning. This fusion might enhance the accuracy and robustness of neural signal decoding, especially when dealing with complex spatial-temporal dynamics. This hybrid technique is already widely used in business data-mining, ”majority-voting” algorithms, and is yet to be incorporated into BCIs.
2) **Adaptive and Online-Learning HMMs:** EEG signals and user brain patterns vary over time (due to fatigue, electrode shifts, learning, etc.), so static HMM parameters can degrade. A key direction is making HMMs adaptive and trainable online. HMMs inherently support incremental inference: as each new EEG sample arrives, the hidden-state probabilities can be updated without restarting the model. Rose et al. [1] explicitly note that to handle nonstationarity one should “update the HMM on-the-fly” during online use. In practice, this means continually adjusting the transition and emission probabilities. For example, one can use an online Expectation–Maximization (EM) algorithm: rather than re-training on batch data, the updates can be applied incrementally as new data streams in (with decaying learning rates). This allows the HMM to gradually adapt to changing EEG distributions. Rose et al. point out that such online updates “are generally unproblematic for HMM training”, suggesting HMMs can naturally incorporate sequential adaptation.
3) **Enhanced Pre-processing Strategies:** Future work can benefit from refining pre-processing techniques—such as advanced noise reduction, artifact removal, and optimized feature extraction procedures—to improve the quality of observable states fed into the HMM. A focus on robust signal-cleaning protocols will reduce ambiguity within state representations.
4) **Interpretability and Explainability:** Further investigation into the interpretative aspects of HMMs can help demystify the relationship between shifts in brain state transitions and specific cognitive functions. Developing visualization techniques and explainable models may foster broader adoption of HMM-based approaches in both clinical and research contexts.
5) **Real-Time HMM-Based BCI Systems:** Applying HMMs in online BCI systems requires meeting strict real-time constraints (low latency, continuous feedback). HMMs have properties well-suited to this: they naturally produce a posterior over hidden states as data streams in, allowing incremental decoding. In practice, an online HMM decoder runs a forward or Viterbi update with each new EEG sample (or short window), outputting the most likely current state. Rose et al. emphasize that unlike a static classifier, the HMM “does not start from scratch” on each new sample – it simply updates its current prediction with incoming data. This means HMM-based BCIs can provide prompt updates. Indeed, in their implementation, the HMM achieved a classification rate of 285 Hz when classifying after each sample, indicating that even with a non-optimized Matlab code the system could process EEG in real time (modern optimized code or hardware could reach far higher rates, as used in speech recognition). Several BCI systems have tested HMMs online. Rose et al. (2013) directly compared HMM and SVM decoders for finger-movement ECoG. Both achieved approximately 90% accuracy, and the HMM had the advantage of naturally updating per-sample; the authors conclude that “HMMs and their characteristics are promising for efficient online brain–computer interfaces. In another example, a hierarchical HMM controlled six degrees of freedom (hand, elbow, shoulder) with performance gains over a conventional HMM. These studies confirm that HMM-based decoders can operate within the timing demands of BCI control.
6) **Specialized Variants:** To address some limitations, researchers have developed variants of HMM. For example, Hierarchical HMMs (HHMM) introduce sub-models for subsets of classes to reduce complexity. Indeed, one ECoG

study found a two-layer hierarchical HMM (H2M2) yielded higher accuracy and lower latency than a flat HMM, indicating a promising area of research & experimentation.

By addressing these future directions, the potential of HMMs in BCI applications can be more fully realized, paving the way for more intuitive, accurate, and adaptive interfaces that translate brain activity into meaningful actions.

REFERENCES

[1] George Rose et al. "Hidden Markov Model and Support Vector Machine based decoding of finger movements using Electrocorticography". In: *National Center for Biotechnology Information* (2013). DOI: 10.1088/1741-2560/10/5/056020. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC3901317/.

[2] Abdalla et al. Maiseli. "Brain–computer interface: trend, challenges, and threats". In: *Brain Informatics* (2023). URL: https://rdcu.be/ekgIw.

[3] Janis Peksa & Dmytro Mamchur. "State-of-The-Art on Brain-Computer Interface Technology". In: *MDPI* (2023). DOI: 10.3390/s23136001. URL: https://www.mdpi.com/1424-8220/23/13/6001.