Application of frequent episode discovery for analyzing multi-neuron spike train data

Debprakash Patnaik\(^1\), P.S.Sastry\(^2\), and K.P.Unnikrishnan\(^3\)

\(^1\)Virginia Tech, \(^2\)Indian Institute of Science, \(^3\)General Motors R&D

Inferring useful information regarding coordinated behavior of a group of neurons from simultaneously recorded spike train data is a challenging problem of current interest in neuroscience. Due to the availability of large sets of such data, there is a lot of interest in developing novel data-analysis techniques that are efficient for this problem \([1]\). Many of the current techniques such as cross correlograms, JPSTH, Principal component analysis etc. becomes computationally very intensive when one wishes to infer relationships among more than a few (typically 3) neurons. The field of temporal data mining in Computer Science is concerned with developing efficient algorithms for analyzing extremely large databases of symbolic sequential data streams \([2]\). In this paper we present that techniques from temporal data mining are very effective for getting information regarding the microcircuits underlying a neural tissue by analyzing the multi-neuron spike train data. Specifically we consider the framework of frequent episodes in temporal data mining. In this framework, the data is viewed as a sequence of (time-ordered) events where each event is characterized by an event type (which is a symbol from a finite alphabet) and a time of occurrence. For spike train data, the event types would be neuron labels and event times would be the time of occurrence of the respective action potentials. The objective is to discover repeating temporal patterns called episodes. A serial episode is an ordered sequence of event types and a parallel episode is an unordered collection of event types. An episode is said to occur in a slice of data if we can find events in the data with appropriate event types that satisfy the ordering constraint of the episode. For example, a serial episode $A \rightarrow B \rightarrow C$ occurs in the data if we can find an $A$ and sometime later a $B$ and so on. There can be other intervening events in between. In the spike train data, if there are strong excitatory connections from $A$ to $B$ and $B$ to $C$, then we can expect to see many occurrences of such a serial episode. Since there are other neurons also spiking, there can be other events in between those corresponding to an episode occurrence. Similarly, a parallel episode $(ABC)$ would occur frequently in the data if the corresponding neurons often fire synchronously. To capture many such types of temporal patterns properly, we extend the frequent episodes framework so that one can impose some temporal constraints on the occurrences of the episodes that are counted \([3]\). These constraints allow us to take care of, e.g., synaptic delays. In this paper we show that this extended episodes framework is well suited for discovering many patterns of interest in spike train data such as order, synchrony and synfire chains. We present algorithms that can efficiently discover serial and parallel episodes (under temporal constraints) that occur sufficiently often in the data. We show through simulations on synthetic and real data that these algorithms are very effective in unearthing interesting temporal patterns in the spike train data involving large number of neurons.

References

