Bio-Inspired Temporal-Decoding Network Topologies for the Accurate Recognition of Spike Patterns

Gianluca Susi\textsuperscript{1,2}
\textsuperscript{1}Department of Electronic Engineering, \textsuperscript{2}Department of Civil Engineering and Computer Science
University of Rome "Tor Vergata", Italy.
gianluca.susi@uniroma2.it

ABSTRACT

In this paper will be presented simple and effective temporal-decoding network topologies, based on a neuron model similar to the classic Leaky Integrate-and-Fire, but including the spike latency effect, a neuron property able to take into account that the firing of a given neuron is not instantaneous, but it occurs after a continuous-time delay depending on the inner state. These structures are able to detect spike sequences composed of pulses belonging to neuron ensembles, exploiting basic biological neuron mechanisms. According to the biological counterpart, with these structures is possible to achieve a high temporal accuracy, but also deal with the natural variability present in spike trains. In addition, the connection of these neural structures at a higher level make possible to afford some pattern recognition problems, operating a distributed and parallel input data processing.

Keywords: Temporal coding; Neuronal modelling; Spiking Neural Network (SNN); Latency; Pattern recognition; Classification; Coincidence detection.

1 Introduction

In sensory systems the recognition of relevant stimuli is made possible by the detection of spike patterns during the processing of peripheral inputs. However, specific neural mechanisms that enable recognition of such sequences are currently unknown [1].

‘Decoding’ refers to the problem of how to ‘read out’ the information contained in a set of neural spike trains and has both theoretical and practical implications for the study of neural coding [2, 3]. Some works highlight that neurons represent information through their mean rates of action potential firing [4], on the other hand data suggest that the temporal structure of spike trains take advantage of reliable and precise latencies [5, 6]. In [7] is shown how time coding and rate coding tend to become confounded, at very short time scales (10 ms or less), and any form of temporal code can be described in terms of very rapid changes in firing rate [8].

Temporal spike coding, together with temporal spike learning, has been revealed in the auditory- and visual information processing in the brain as well as in motor control [9], and its use in neuro-prosthetics is essential along with applications for a fast, real-time recognition and control of sequence of related processes [10]. In [7] is highlighted the relevance of considering the timing of spikes across a population of afferent neurons.
In this paper, very simple and effective temporal-decoding network topologies will be presented, based on the spike latency effect, an intrinsically continuous neuron property able to take into account that the firing of a given neuron is not instantaneous, but occurs after a continuous-time delay [11], depending on the inner state. With these structures is possible to process spike trains arisen by neuron ensembles. Also considering other bio-plausible phenomena, in the next sections a method for the tuning of such structures will be presented.

2 Neuron and network model

The structures presented in this paper are based on the Leaky Integrate and Fire with Latency neuron model [12, 13]. This kind of neuron is characterized by a real non-negative inner state ($S$), and is able to work in two modes of operation, depending on whether $S$ is above or below a certain value; this value, called firing threshold ($S_{th}$), is equal to $1 + d$, where $d$ is a positive value called threshold constant.

Passive mode is the operating mode of the neuron when its inner state is less than $S_{th}$ ($S < S_{th}$); otherwise, the neuron is said to be in active mode ($S \geq S_{th}$). In the passive mode the neuron is a simple integrator, characterized by a linear subthreshold decay with constant $Ld$, such that in a time interval between two consecutive input spikes, $\Delta t$, the inner state decreases of a quantity $T_f = Ld \Delta t$. Conversely, in the active mode the neuron is ready to generate a spike. However, the output spike is not produced immediately, but after the latency time, called in the model time-to-fire ($t_f$). The inner state and the time-to-fire are related through the following bijective relation, called firing equation

$$t_f = \frac{1}{S - 1}$$

As time passes, the time to fire decreases, and inner state grows up, but it remains sensitive to possible excitatory or inhibitory input spikes. For $S = S_{th}$ (i.e., the minimum inner state possible in active mode) the time-to-fire is equal to $t_{f, max} = 1/d$ (i.e., the maximum time-to-fire).

Once the $t_f = 0$, the neuron definitively generates a spike. The amplitude of each spike generated by the network is characterized by the product $P_j = P_r \cdot P_w(x,j) = P(x,j)$. The quantity $P_w(x,j)$ (or equivalently $P_{wij}$ if the target neuron is implicitly indicated) is the postsynaptic weight, that characterizes the connection strength from the neuron $j$ to a certain neuron $x$. The quantity $P_r$, is the presynaptic weight, that represents the value of the spike amplitude as it has been generated by the firing neuron.

Once the neuron generates a spike, its state is reset to its resting potential (conventionally, $S = 0$) for a time defined as the absolute refractory period ($t_{arp}$).

The inner state evolution of a neuron in response to a spike sequence is illustrated in Fig. 1, where two incoming spikes modify the initial inner state $S_{po}$ (i.e., previous-state-0) of a target neuron.

For the purpose of emulating a continuous-time behavior, an event-driven approach is required, as described in [13], where the network model is characterized by the following global parameters: $P_r, Ld, d$ (and then, $S_{th}$); in the present work the same global parameters are considered.

For the sake of simplicity, in the structures presented in this paper zero-delays axonal connections are considered, and synaptic plasticity phenomena are not taken in account. However, the relations that will be obtained below remain compatible with such cases.
Figure 1. An example of the inner state behavior of a neuron in both passive and active modes is illustrated. An incoming excitatory pulse at $t_1$ causes an instantaneous increase of the state from $S_{p0}$ to $S_{p0} + P_r P_w$. When, at $t_2$, a second excitatory pulse is applied, the state increase his value from $S_{p2}$ to $S_{p2} + P_r P_w$. The firing is not instantaneous but occurs after $t_f$. Finally, after the firing, the neuron is reset to its resting potential for a time equal to $t_{arp}$ [13].

3 Dynamic summation in bio-plausible neuronal structures

A spike sequence can be quite exhaustively defined by amplitudes and arrival times of its components. Many attempts described in scientific literature are pointed to perform the recognition of specific spike sequences through bio-plausible neural structures (e.g., [14-16]). Typically, the recognition process is signaled by the activation of an indicator unit (e.g., a single neuron, a particular combination of neurons, or a whole network); with the aim of simplifying the analysis, in this paper a single neuron (called target neuron, TN) is used as indicator unit, so that its firing identifies that the recognition of the sought input spike sequence has occurred.

The purpose of this section is to introduce a key-mechanism of the neuron, here indicated as dynamic summation. Such mechanism can be exploited for the realization of many simple structures whose operation takes in account precise spike times of the input sequence. In particular, basic aspects of such mechanism will be explained in the next subsection.

3.1 Target neuron and working mode activity level

The task of a TN is to generate an output spike when a number of contributions is received in a certain time interval (Fig. 2). Actually, the target neuron becomes active when the following two conditions are simultaneously satisfied:

- the sum of the incoming spike amplitudes is greater than $S_{th}$;
- the input spikes are properly synchronised.

For an exhaustive explanation of the subthreshold decay effect and the interplay between pulse amplitudes and arrival times characterizing the spike sequence, the reader may consult [17].

With reference to Fig. 2, in order for the TN to generate a spike, the following relation has to be satisfied:

$$P_r (P_{wa} + P_{wb} + P_{wc}) - L d (\Delta t_1 + \Delta t_2) > S_{th}$$

where $P_r$ is the pre-synaptic weight, $P_{wa}, P_{wb}$ and $P_{wc}$ are the post-synaptic weights of the branches afferent to TN and $\Delta t_1, \Delta t_2$ are the interspike intervals (ISIs) characterizing the specific spike sequence.
The working mode activity level (WMAL) is a parameter that explicates the number of input spikes, simultaneous and of equal amplitude, necessary to lead the inner state of TN over the threshold, starting from the resting potential (Fig. 3). Of course, the concept is related to a condition upon the weights of the afferent branches; it is particularly useful to consider the WMAL for the synthesis of structures oriented to the recognition of spike sequences.

In the design phase, hypothesizing the same value for the weights of the afferent branches, the WMAL is obtained taking into account the weight $P_j = p_i p_{wj}$ (identical for the neurons afferent to the TN) and the threshold value $S_{th}$:

- to obtain $WMAL = 2$, it must be imposed
  \[ \frac{S_0}{2} < P_j < S_0 \]  \( (3) \)
- to obtain $WMAL = 3$, it must be imposed
  \[ \frac{S_0}{3} < P_j < \frac{S_0}{2} \]  \( (4) \)
- to obtain $WMAL = n$, it must be imposed
  \[ \frac{S_0}{n} < P_j < \frac{S_0}{n-1} \]  \( (5) \)
Fig. 3. Case of WMAL = 3. In reference to the architecture shown in Fig. 2, three identical and simultaneous pulses ($P_{wa} = P_{wb} = P_{wc}$; $t_a = t_b = t_c$) are necessary to let TN spike. A smaller amplitude or number of input spikes, or a smaller synchronism among them, could make the TN not able to generate the output spike.

### 3.2 Time-amplitude uncertainty

The inner state of a neuron at a time $t_0$ does not provide information on the previous neuron activity: even though the initial state of the neuron is known, after a certain interval of activity it can reach a given inner state in infinite manners.

In this paper will be presented bio-inspired structures for the recognition of spike patterns. In the tuning phase it must be considered that the components of the input spike-sequences can present timing errors, or amplitudes not equal to the expected ones. With the aim of analyzing in detail this issue, in this paragraph will be shown the time-amplitude uncertainty phenomenon.

Useful relations for the design phase of detector structures will be given considering the internal state of the neuron as a function of its previous state, both for passive and active modes (i.e., respectively, subthreshold and overthreshold), hypothesizing a temporal neighbourhood in which the neuron does not switch operation mode (i.e., no threshold crossing).

#### 3.2.1 Subthreshold uncertainty

Considering a neuron in passive mode is straightforward to express the inner state as a function of its previous state. Said

$$S_1 = S(t_i)$$

and

$$S_2 = S(t_2)$$

it can be written

$$S_2 = S_1 - K_d(t_2 - t_1)$$
Figure 4: Subthreshold uncertainty phenomenon: assuming that at time $t_2$ the internal state of the neuron is $S_1$, there is uncertainty on the activity that has lead the system on this state. With the aim of highlighting such uncertainty, a compatible subthreshold curve is plotted (in red).

3.2.2 Overthreshold uncertainty

Also considering a neuron in active mode it is possible to express its inner state as a function of its previous state. Said

$$S_1 = S(t_1)$$
$$S_2 = S(t_2)$$

we obtain

$$t_{f1} = \frac{1}{S_1 - 1}$$
$$t_{f2} = \frac{1}{S_2 - 1}$$

where

$$t_{f2} = t_{f1} - (t_2 - t_1) = t_{f1} - t_2 + t_1 = \frac{1}{S_1 - 1} - t_2 + t_1$$

therefore

$$S_2 = 1 + \frac{1}{t_{f2}} = 1 + \frac{1}{\frac{1}{S_1 - 1} - t_2 + t_1}$$

Figure 5: Overthreshold uncertainty phenomenon: assuming that at time $t_2$ the internal state of the neuron is $S_2$, there is uncertainty on the activity that has generated such a situation. With the aim of highlighting such uncertainty, a compatible overthreshold curve is plotted (in red).

During the design of the structures proposed in the following sections, it will not be always needed to specify both all the amplitudes of the input pulses, and all the ISIs of the spike sequence. Nevertheless, with the aim of unambiguously identify a sequence, it may be useful to refer to the relations obtained above.
4 Spike-timing sequence detectors

Spike-timing sequence detectors (STSD), are classes of structures able to reveal specific ISIs present in input sequences. These structures provide also the opportunity to set the tolerance of the detection in respect to the spike timings of the sequences to recognize.

Different kinds of realization are possible for these structures: they may differ in the number of spike trains (i.e., input lines) covered, number and type of neurons involved, or other aspects, like the number of neuronal phenomena exploited in the computation. In the course of this discussion, an order of presentation based on their complexity of operation will be followed.

In the structures presented in this paper, spike sequences composed of pulses belonging to different neurons will be considered, as parallel spike trains coming from a neural ensemble [19, 20, 16]. By operating on some of the parameters presented above it is possible to tune the structures on particular intervals, in order to detect specific sequences.

The relations between the times of the input spikes arrival and the effective firing of the target neuron will be shown by specifying the activation domain, and the operation of the presented structures will be shown by means of temporal diagrams.

4.1 Activation domains

In the following sections are presented different classes of structures able to detect specific ISIs between couples of pulses of an input sequence. Brain regions involved in object recognition in general must deal with the natural variability present in spike trains [1]. To take account of this necessity of flexibility, with the aim of maintain a bio-plausible behaviour, in the class of structures presented in this paper, a certain tolerance upon the arrival times of the input pulses can be set. Fig. 6 illustrate different kind of activation domains, namely the variation ranges of the input spikes, for which the structure is still able to detect the input sequence and to activate the TN.

Figure 6: Different activation domain classes: single (a,b) and multiple (c). Green lines represent the tolerance ranges on the timings of the input spikes (orange and blue).

Note that, with the aim of facilitating the approach, in this paper the arrival time of the pulse from \( EI_1 \) will always considered preceding that of \( EI_2 \) (i.e., \( t_2 > t_1 \)), as shown schematically in Fig. 6.

4.2 Direct STSD

Direct STSD is the simplest class of structures able to reveal specific ISIs present in input sequences.

The properties of these structures strictly depend on the global parameters \( (P_r, Ld, d) \), and on the value of \( P_r \), defined as the common post-synaptic weight of the synapses afferent to the TN.
Although a greater number of branches permits the detection of spike sequences of higher cardinality, as a basic example it will be analyzed the **Simple direct STSD**, that consists of two branches connected to the external inputs \((E_1, E_2)\) and afferent to a \(TN\). This structure is able to operate with spike sequences of cardinality 2. For the sake of simplicity, for this structure unitary amplitudes will be considered for the input spikes for a better understanding of the operation principle.

The topology of this particular architecture is shown in Fig. 7

![Figure 7: Topology of the Simple direct STSD structure. The detection is positive when the neuron TN produces a spike; otherwise the detection is negative.](image)

In Fig. 7, \(P(TN, E_i)\) represents the product between pre- and post-synaptic weights concerning the synapse \(E_i-TN\), and \(P(TN, E_j)\) represents the same kind of product concerning the synapse \(E_j-TN\). With a proper choice of the parameters, the system is able to work with \(WMAL=2\). This means that \(TN\) could become active in presence of two input pulses from \(E_1\) and \(E_2\). For example, if \(P(TN, E_1) = P(TN, E_2) = 0.8\), and the spikes from \(E_1\) and \(E_2\) arrive in a very close succession, \(TN\) becomes active with state equals to:

\[
S(TN) \approx 1.6 \ (> S_{th}) \tag{15}
\]

such that an output spike is produced, after a latency interval \(t_f\) defined by the firing equation (1).

In this case, the output spike generated by \(TN\) reveals that the structure has received in its input a sequence composed of two spikes close enough. On the other hand, if the arrival times of \(E_1\) and \(E_2\) (respectively \(t_1\) and \(t_2\)) are far, only the value \(S(TN) = 0.8\) is present at \(t_2\). In facts, for the **subthreshold decay** phenomenon, the inner state of \(TN\) at time \(t_2\) falls to

\[
S(TN) = 0.8 - (t_2 - t_1) Ld \tag{16}
\]

and the new incoming spike from \(E_2\) is not able to make \(TN\) active; in this case, the detection is negative. Whereas, the detection can be positive if the difference \((t_2 - t_1)\) is smaller.

![Figure 8: A two-input direct STSD (i.e. Simple direct STSD); external inputs are affected sequentially by the spikes of the proposed sequence.](image)
Temporal diagram (Fig. 9, a and b) and activation domain (fig. 10) of the Simple direct STSD are plotted below, considering $P_r = 1$, and $P_w (TN, EI_1) = P_w (TN, EI_2) = 0.8$.

![Temporal diagram and activation domain](image)

**Figure 9**: Temporal diagram of the Simple direct STSD. As depicted in a), if the second spike arrives in the temporal range between the first spike and the value $t_o$, the inner state of $TN$ exceeds the threshold, then an output spike is generated. If the time of the second spike exceeds $t_o$, as depicted in Fig. 9b, the detection becomes negative due to the subthreshold decay phenomenon.

**Figure 10**: Activation domain of the Simple direct STSD: for sequences characterized by interspike intervals comprised between 0 and a certain value $t_o$, the detection is positive; otherwise the detection is negative.

It is possible to tune the structure with the aim of detecting a particular ISI range. The following two relations represent a possibility to properly set the structure parameters to make it sensitive to a particular interval ($t_2 - t_1$), modifying the values $P (TN, EI_1)$ and $P (TN, EI_2)$.

The first condition allow the TN to exceed the threshold

$$P_r P_t - (t_2 - t_1) L d > 1 + d$$

and the second condition imposes a $WMAL = 2$

$$\frac{1}{2} \frac{1+d}{P_r} < P_t < \frac{1+d}{P_r}.$$  

Of course it is possible to modify the structure in order to recognize sequences of cardinality greater than 2. For this end, in the following sections some considerations will be presented.

### 4.3 Delayed STSD

This class of structures is characterized by the presence of one or more delay neurons (DNs), i.e. simple excitatory neurons in the middle of a branch, able to create a transmission delay on the connection they are inserted, and capable of obtain useful effects for the computation.

The operation is based on the fact that the pulses coming from the two branches are able to generate a spike in TN only if they are properly distant in time; in this scenario, the delay neuron provides to slow down the first spike so that it can converge to the TN temporally closer to the second input.
The properties of these structures strictly depend on the value of Pwt (i.e. the common post-synaptic weight of the synapses afferent to the TN) and on the values of Pwd (i.e. the post-synaptic weights of the synapses afferent to the DNs), in addition to the global parameters.

### 4.3.1 Simple delayed STSD

The previous discussion on the Simple direct STSD, oriented to spike sequences of cardinality 2, can be easily extended to define structures that support more complex activation domains. If the topology is modified as in fig. 11, that is, characterized by the presence of one DN, the activation domain may be shifted of a time equal to \( dt \) (i.e. the latency time provided by DN) with respect to the arrival time of the first spike. In this section, the upper branch will be indicated as the delay branch.

![Figure 11: Topology of the Simple delayed STSD.](image)

The value of \( dt \) can be modified by varying the values of the weights. Of course, this neuron must have a \( WMAL = 1 \).

Temporal diagram and activation domain of the Simple delayed STSD are reported in Fig. 12 and Fig. 13. By analyzing Fig. 11, it is easy to deduce the requirements for the correct operation of the structure: the activation of DN requires \( WMAL = 1 \), whereas \( WMAL = 2 \) is needed for the activation of TN.

![Figure 12: Temporal diagram of the Simple delayed STSD. The structure is able to exhibit an output spike only for a certain range of ISI, characterizing the cardinality-2 sequence.](image)
In addition, with the aim of guaranteeing the activation of the neurons of the input layer, the following relations have to be satisfied

\[ E_{I_1} \& E_{I_2} > 1 + d \]  

For DN we have

\[ P(DN,E_{I_1}) > 1 + d \]  

namely

\[ P_r P_w(DN,E_{I_1}) > 1 + d \]  

that is to say

\[ P_w(DN,E_{I_1}) > \frac{1+d}{P_r} \]  

On the other hand, to guarantee the output spike, assuming the favorable simultaneity condition on the target, the following relation have to be satisfied

\[ P_r [P_w(TN,DN) + P_w(TN,E_{I_2})] > 1 + d \]  

then

\[ P_w(TN,DN) + P_w(TN,E_{I_2}) > \frac{1+d}{P_r} \]  

Indicating with \( \Delta t_{in} \) the time difference between input pulses and \( \Delta t_{out} \) the time difference between the pulses afferent to the target, it is possible to merge the previous conditions, as follows:

\[ P_r [P_w(TN,E_{I_2}) + P_w(TN,DN)] - Ld \left| \frac{1}{(ext_{2-1})} - \frac{1}{(ext_{1-1})} - \frac{1}{[P_r P_w(DN,E_{I_1})-1]} + \Delta t_{in} \right| > (1 + d) \]  

This inequality permits to identify relations that allow the design of such structures for specific purposes.

However, this form is incompatible with an univocal design of the structure because an excessive number of degrees of freedom is present. In order to achieve the simultaneity condition of the two contributes on \( TN \), the following equality must be verified:

\[ t_f(E_{I_1}) + t_f(DN) = \Delta t_{in} + \Delta t_{out} + t_f(E_{I_2}) \]  

with \( \Delta t_{out} = 0 \). Then

\[ t_f(E_{I_1}) + t_f(E_{I_2}) = \Delta t_{in} + t_f(E_{I_2}) \]  

It can be noted that for spike sequences in which pulses are far away, such a structure would be useless. In that case is necessary to use more than one \( DN \) in cascade, in order to satisfy the following relation:

\[ t_f(E_{I_1}) + t_f(DN_1) + t_f(DN_2) + \cdots + t_f(DN_n) = \Delta t_{in} + t_f(E_{I_2}) \]
In addition, in order to recognize sequences of greater cardinality, it is necessary to increase the number of delay branches, such that the contributions can arrive simultaneously to the target neuron, which must be of type WMAL n.

In relation to the Simple delayed STSD, let us analyze the tuning procedure.

Indicating Tol the permissible error on the arrival time of the second pulse still capable to let the TN spike, and MaxTol the maximum value of Tol that can be used, the tuning of the structure can begin with the computation of the maximum tolerance (MaxTol) supported by the Ld considered; under proper considerations, it can be written

\[
Ld = \frac{(2P_r P_T - 1 - d)}{Tol}.
\]

Placing \(P_t\) to the upper end, we obtain

\[
Ld = \frac{\left\{2P_r \frac{1+d}{P_T} - 1-d\right\}}{\text{MaxTol}}
\]

where

\[
\text{MaxTol} = \frac{2(1+d)-1-d}{Ld} = \frac{(1+d)}{Ld}.
\]

At this point, the requirements necessary for the operation of the structure imposes the relations (19) and (20). With the aim of confining the degrees of freedom of the structure it is possible to proceed as follows. The WMAL of TN can be subdivided evenly with respect to the number of afferences (i.e., make equal the post-weight of the neurons afferent to TN)

\[
P_r P_w(TN, EI_2) = P_r P_w(TN, DN) = P_r P_t.
\]

In order to assure the output spike, from (4) we obtain

\[
2P_r P_t > (1 + d)
\]

whereas, in order to guarantee that the target does not work in WMAL = 1, must be

\[
P_r P_t < (1 + d).
\]

Merging the last two conditions, for a two-input Simple delayed STSD we obtain

\[
1/2 [ (1+d)/P_r] < P_t < [(1+d)/P_r]
\]

In this way, imposing the simultaneity condition on TN, an output spike will be certainly produced and the target will work at a WMAL = 2.

In reference to the interpulse interval between the two input spikes, the more \(P_t\) tends to \([(1+d)/P_r]\), the more the structure is tolerant to temporal errors of the second input pulse; conversely, the more \(P_t\) tends to \([(1+d)/2P_r]\), the less the structure is tolerant to temporal errors of the second input pulse.

### 5 STSD-based temporal coding classifiers

Using the elementary structures presented in the previous sections, some pattern recognition problems can be easily afforded.

In an-dimensional space, a single object can be represented as a set of n-reference features. Each axis is referred to a time interval that permits to quantify a reference feature, so that objects belonging to different classes will
be mapped in a n-dimensional "timing space" [17]. An operation scheme using Simple delayed STSD is depicted in Fig. 14.

6 Conclusion

In this work have been presented some bio-inspired temporal-decoding network topologies based on a bio-plausible neuron model, for the accurate recognition of spike patterns. These structures are able to detect specific input spike sequences in a continuous time domain. The connection of these neural structures at a higher level, make possible to process the information operating a distributed and parallel input data processing.

Future developments will include the application of the spike-timing dependent plasticity (STDP) algorithm with the purpose of auto-setting tuning parameters as the postsynaptic weights.

Figure 14. Scheme of a STSD-based temporal coding classifier. Case of three features, each one characterized by means of three classes: a) class 1 recognizer set; b) class 2 recognizer set; c) class 3 recognizer set; d) single class domains, in the three dimensional feature space [17].

REFERENCES


