Written by Yarden Nativ, mail: nativyardenac@gmail.com

5th Recitation 27.4.22

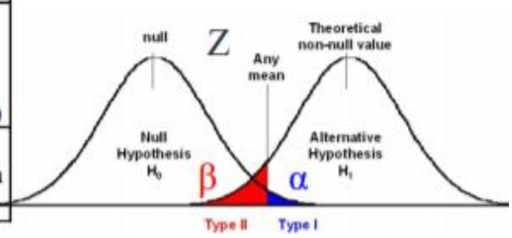
ROC, Quiz rehearsal

Discrimination

Discrimination tools evolved to solve the following problem: Given a set of measurements for a neuron, for example $\{r_0, r_1, \ldots, r_i\}$ and a set of stimuli $\{H_0, H_1, \ldots, H_i\}$, we look for a function to relate between the elements in each set. Practically, discrimination is commonly used to define how good can the neuron separate between different given stimuli.

Simplifying the problem to two stimuli, we can assume that for a given threshold of firing rate, Z, if $r \geq Z$ than the corresponding stimulus is H_1, otherwise it's H_0. This assumption is illustrated by the following graph and table:

		Actual value	
		P	N
Prediction	P	* True P * Detection * 1-β(z)	* False P * False detection /False alarm * Type I error α(z)
	N	* False negative * Miss detection * Type II error β(z)	* True N * Correct rejection * 1- α(z)



Basic terms in discrimination:

- $\alpha(Z) = P(r \geq Z|H_0)\ \ false\ positive$
- $\beta(Z) = P(r < Z|H_1)\ \ false\ negative$
- $sensitivity = 1 - \beta(Z)$

  $sensitivity = 1 \rightarrow recognizes\ all\ positives.$
- $specificity = 1 - \alpha(Z).$

  $specificity = 1 \rightarrow recognizes\ all\ negatives.$
- $Power\ of\ the\ test\ 1 - \beta$

Written by Yarden Nativ, mail: nativyardenac@gmail.com

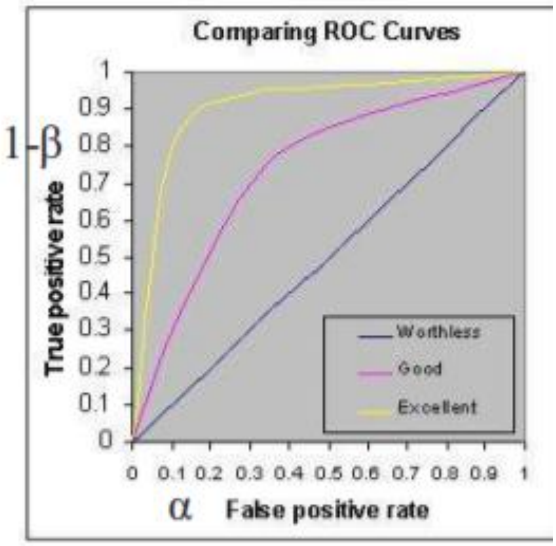
ROC:

Practically, α errors are considered more problematic (for example in drug research or radars), so we prefer to reduce maximally the α. Receiver Operating Characteristic/ Curve (ROC) is a graphic tool to represent the ratio between False Positive (FP) and True Positive (TP) for different values of Z. How to build it:

1. For a given Z, we calculate the TP (1-β) and FP (α).
2. For the calculated TP and FP we assign a dot in the ROC curve.
3. Evaluate the classification with the integral $\int(1-\beta)d\alpha$

Class Discussion:

- What does an integral of $\frac{1}{2}$ stands for? And 1?
- Explain the claims of the following graph:



Examples from past exams:

2007 exam: Given the following probabilities of evoked potential amplitudes:

Neuron rate (spikes/sec)	0	10	20	30	40
Prey	0	0.2	0.2	0.3	0.3
Predator	0.3	0.3	0.2	0.2	0

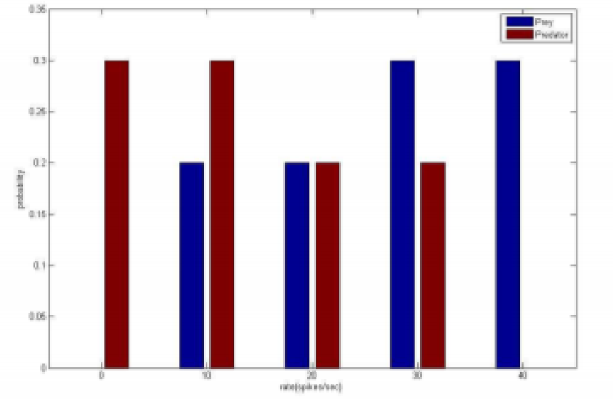
The behavior of the animal may be characterized by:

$$P(TP) = \sqrt{P(FP)}$$

- Plot the ROC curves of the behavior and neuronal discrimination.
- Is the single neuron sufficient to predict the behavior?
- What is p[success] for the two statistics?
  (reminder: P(success) = ∫(1-β)dα)

Solution:

For ROC curve, we need to define first what is H_0 and what is H_1. To do so, we will plot first the probability distributions of each one:



From the distributions, one can assume H_0 is for a predator, and H_1 for prey.

Now we can calculate for a range of Z the FP and TP of the neuron:

$$X\ axis\ of\ ROC: \quad FP(Z) = P(rate \geq Z|H_0)$$

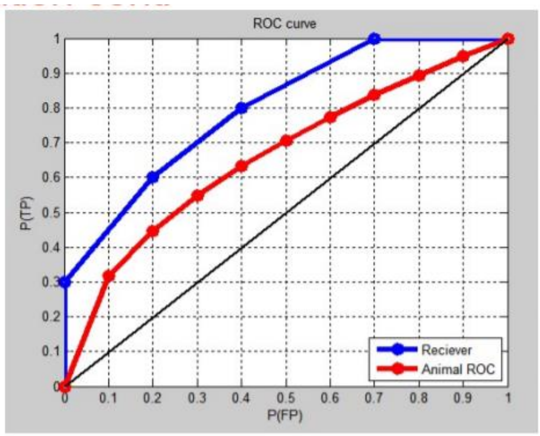
$$Y\ axis\ of\ ROC: \quad TP(Z) = 1 - P(rate < Z|H_1) = P(rate \geq Z|H_1)$$

Z	-10	0	10	20	30	40	50
FP	1	1	0.7	0.4	0.2	0	0
TP	1	1	1	0.8	0.6	0.3	0

To calculate the ROC curve of behavior we can use $TP = \sqrt{FP}$:

FP	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
TP	0	0.3	0.45	0.55	0.65	0.7	0.8	0.85	0.9	0.95	1

And the ROC curves are:



Given the two curves we can understand that a single neuron is more effective than behavior to discriminate between the two stimuli.

To calculate $P(sucess)$ we will use the sums of triangles and trapezoids for the neuron:

$$P(success)_{neuron} = \frac{(0.3+0.6)\cdot 0.2}{2} + \frac{(0.6+0.8)\cdot 0.2}{2} + \frac{(0.8+1)\cdot 0.3}{2} + \frac{(1+1)\cdot 0.3}{2} = 0.8$$

And for the behavior we can use the definition:

$$P(success)_{behavior} = \int_0^1 x^{\frac{1}{2}}dx = \left[\frac{2}{3}x^{\frac{3}{2}}\right]_0^1 = \frac{2}{3}$$

Written by Yarden Nativ, mail: nativyardenac@gmail.com

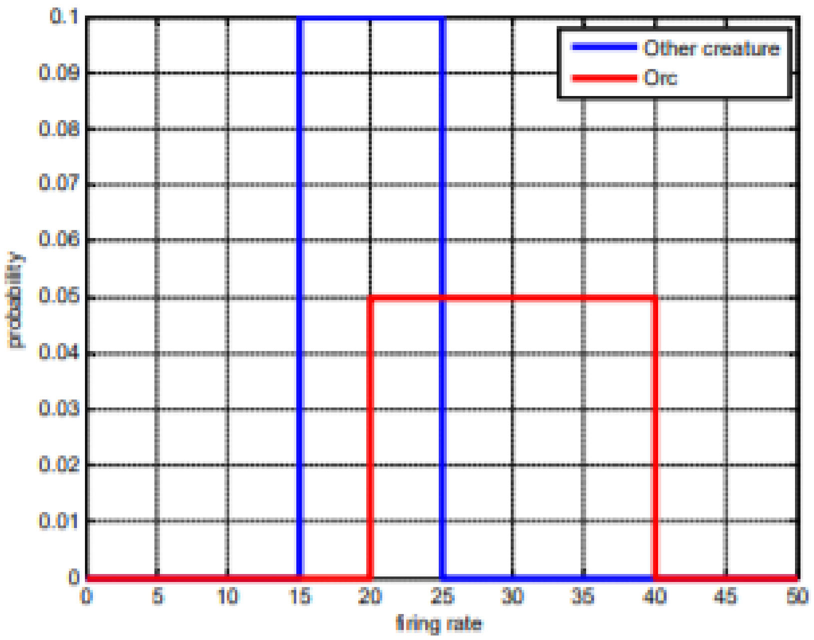
2010 exam: Bilbo Baggins has a neuron which is part of the Orc sensing system. Upon sensing an Orc the neuron's firing rate is described by a uniform distribution in the range [20-40] while upon sensing any other creature the firing rate is taken from a uniform distribution in the range [15-25].

a) Draw the ROC curve of the Orc identification.

b) Calculate the classification performance.

Solution:

In such questions, it is very useful to plot first a histogram of the firing rates, which will help us determine what is H_0 and what is H_1.



Now we can calculate:

$$Z = 10: p(FP) = 1, P(TP) = 1$$

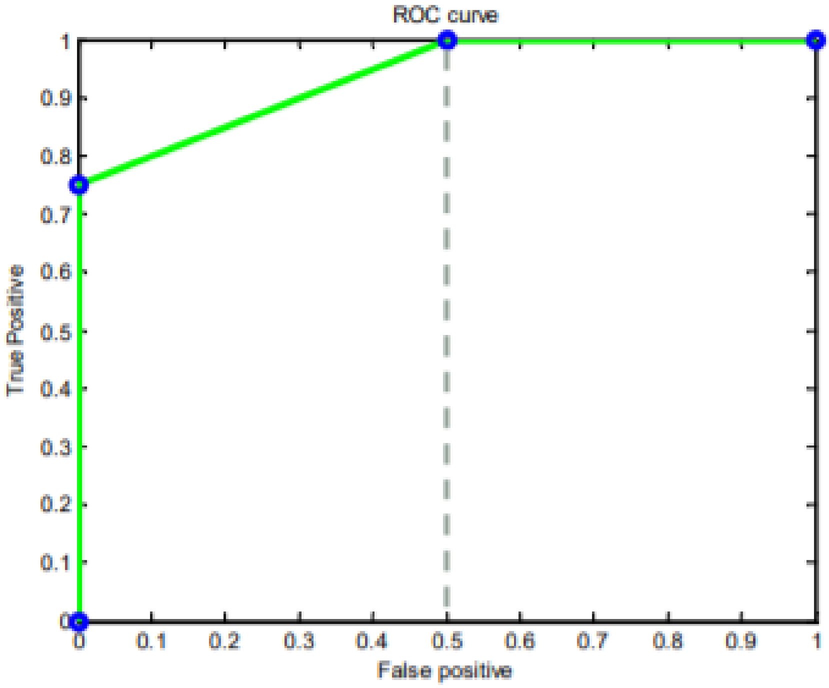
$$Z = 15: p(FP) = 1, p(TP) = 1$$

$$Z = 20: \ p(FP) = 0.5, p(TP) = 1$$

$$Z = 25: p(FP) = 0, p(TP) = 0.75$$

$$Z = 40: p(FP) = 0, p(TP) = 0$$

Therefore, the ROC curve is:



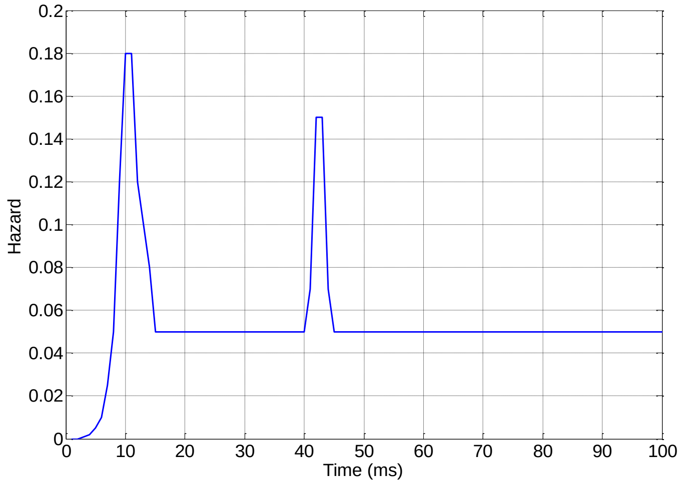
To calculate the performance, we can calculate the areas under the graph using a sum of trapezoid and two rectangles:

$$\int TPdFP = 1 \cdot 0.5 + \frac{0.75 + 1}{2} \cdot 0.5 = 0.9375$$

Written by Yarden Nativ, mail: nativyardenac@gmail.com

Quiz rehearsal: Extracting the autocorrelation from the Hazard function

Given the following hazard function:



A. Sketch the autocorrelation function ±1 second and explain your computations.

B. Is the neuron Poisson, Regular, Bursty? Explain your answer.

Solution:

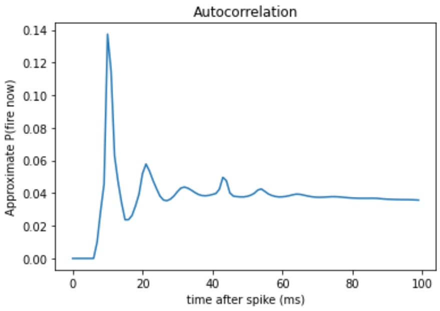
The Hazard function describes the probability of spike firing for a given interval given that no spikes were before. When Hazard function is a straight line we can estimate that this is a Poisson neuron. For example, in the given illustration function, the line is at the level of 0.05 therefore this is a Poisson neuron with a firing rate of 50 spikes per second.

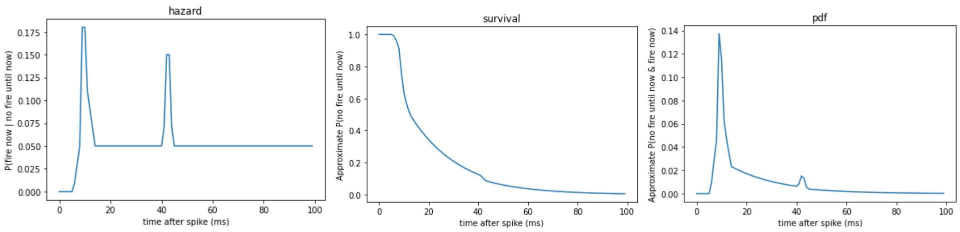
Over the straight line we see few changes from classic Poisson neuron: First, in the first 10 ms there is a refractory period with lower probability for spike. After this refractory period we notice a little horn representing elevation of the probability for spikes in this interval, therefore bursting. A second and lower horn comes later at 42 ms (32 ms after the previous bursting period.

We expect to see both the burstings and the refractory period in the Autocorrelation function, the burstings as two horns in $\tau_1 = 10\ ms$ and $\tau_2 = 42\ ms$, the first one higher than the second. Now we need to add possible combinations between the two times, therefore we expect to see elevation in $\tau = 20ms, 30ms, 40ms, \ldots$ but becoming smaller and smaller. The same about $\tau = 42ms, 84ms, 126ms, \ldots$ and possible combinations between the bursts $\tau = 52ms, 62ms, 72ms, 92ms, 94ms, \ldots$

Written by Yarden Nativ, mail: nativyardenac@gmail.com

Using a simulation we get the following Autocorrelation (Thanks to Benjamin Menashe for simulation):





Written by Yarden Nativ, mail: nativyardenac@gmail.com

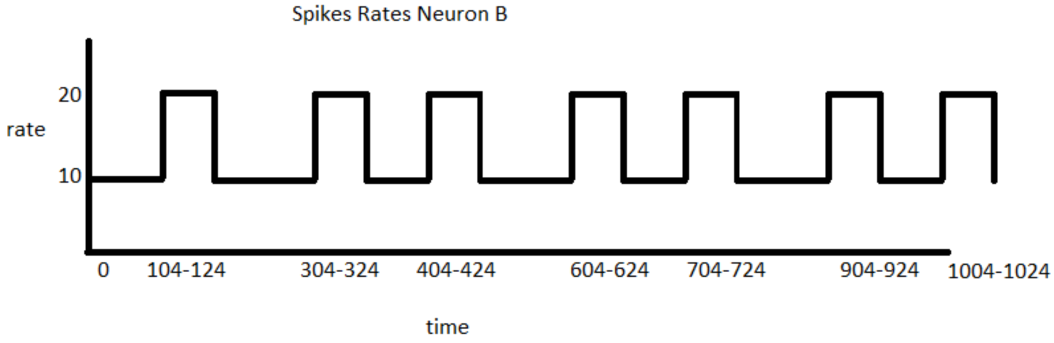
Quiz rehearsal: Autocorrelation for excited Poisson neuron

Neuron A is a regular neuron, alternating between ISIs of 100 and 200ms (e.g. spike at 100, 300, 400, 600, 700, 900, etc.). Neuron B is a Poisson neuron excited by neuron A. Every spike of neuron A leads to an increase in neuron B firing rate from a baseline of 10 spikes/sec to 20 spikes/sec, for a period of 20 ms, following a delay of 4 ms.

Sketch the autocorrelation functions of neuron B in the range of ±500ms, normalize to rate.

Solution:

We will draw first the neuron B:

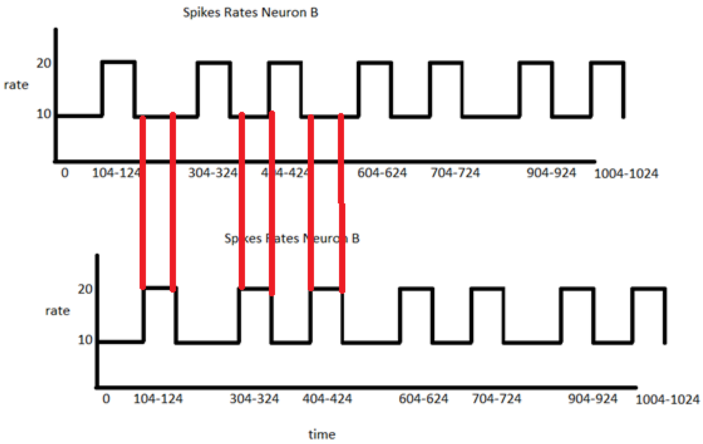


We'll start by focusing on the positive possible time lags by indentation between 1 and 500 ms, so that our strategy is to first count the number of spikes within 500 ms duration under different indentations, and then normalizing to rate. The value of normalization should be $\frac{1}{\frac{1}{1000} \cdot numberOfSpikes} = \frac{1000}{numberOfSpikes}$.

There are 3 saw teeth within 500 ms, therefore the number of spikes is $0.02 \cdot 3 \cdot 20 + 0.01 \cdot (500 - 3 \cdot 20) = 5.6$. Therefore, the normalization value is $\frac{1000}{5.6}$.

We'll calculate the following identations:

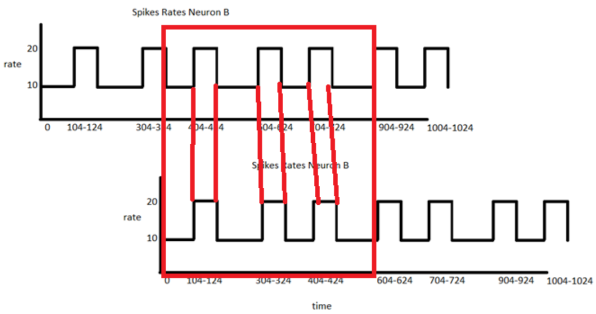
A. After indentation of 20 bins, all the saw teeth will cause full incoordination between all the saw teeth:



In 500 ms duration, the number of spikes which meet each other in such case includes 6 saw teeth of 20 bins where the probability of meeting in each one is $0.01 \cdot 0.02$ and the rest with probability of meeting each other of 0.01^2. Therefore, in this time lag $\tau = 20$ the autocorrelation rate value is:

$$Q_r(\tau = 20) = [0.01 \cdot 0.02 \cdot 6 \cdot 20 + 0.01^2 \cdot (500 - 6 \cdot 20) \cdot \frac{1000}{5.6} \approx 11.08$$

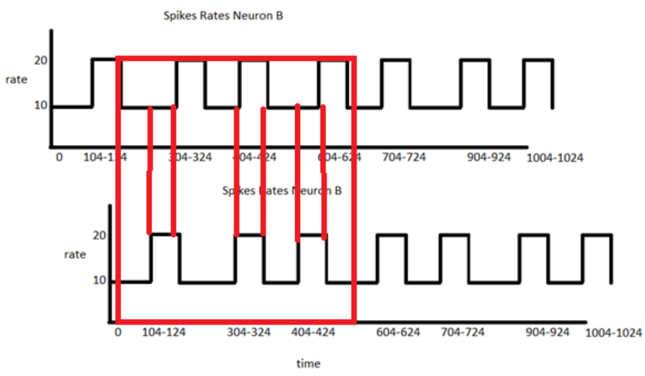
B. After indentation of 300 bins, all the saw teeth will cause full coordination between all the saw teeth:



In such identation there are 3 saw teeth of 20 bins where the probability of spike meeting in each one is 0.02^2, and in the rest the probability of meeting is 0.01^2. Therefore, in this time lag $\tau = 100$, the autocorrelation rate value is:

$$Q_r(\tau = 300) = [0.02^2 \cdot 3 \cdot 20 + 0.01^2 \cdot (500 - 3 \cdot 20) \cdot \frac{1000}{5.6} \approx 12.14$$

C. In the indentation of 100 bins, only one saw tooth will cause full coordination between all the saw teeth:



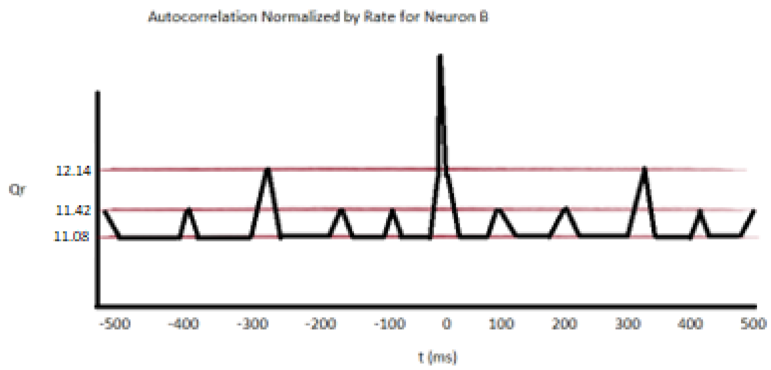
In such identation there is 1 saw tooth of 20 bins where the probability of spike meeting in each one is 0.02^2, 4 saw teeth of 20 bins where the probability of meeting in each one is $0.01 \cdot 0.02$ and the rest with probability of meeting each other of 0.01^2. Therefore, in this time lag $\tau = 100$, the autocorrelation rate value is:

$$Q_r(\tau = 100) = [0.02^2 \cdot 1 \cdot 20 + 0.02 \cdot 0.01 \cdot 4 \cdot 20 + 0.01^2 \cdot (500 - 5 \cdot 20) \cdot \frac{1000}{5.6} \approx 11.42$$

D. Indentations around this three peaks will be calculated based on the following idea for the first indentation of 20 ms. For the indentations $\tau \in ([1,19] \vee [81,119] \vee [181,219] \ldots)$ there are two sub-sequences: the first one is of saw teeth of 20 ms start which start at the bins given by $300n + 4$ and the other one of saw teeth starting with the bins of $300n + 104$.

In the indentation of the subsequence $\tau \in (300n + 4 | [1,19] \vee [281,319])$ there will be three overlap regions between the saw teeth, while in the rest $\tau \in (300n + 104[81,119] \vee [181,219] \vee [381,419])$ only one overlap region. In each overlap region every one bin indentation while cause a change between bins with probability of 0.02^2 and $0.02 \cdot 0.01$, while in the rest of the bins outside the overlap regions the probability will be 0.01^2. Therefore every indentation of one bin will cause a linear change of either increase or decrease in probability between 0.02^2 and $0.01 \cdot 0.02$.

Therefore the autocorrelation function will look like this:



Written by Yarden Nativ, mail: nativyardenac@gmail.com

Quiz rehearsal: Excitatory cross correlation by pacemaker

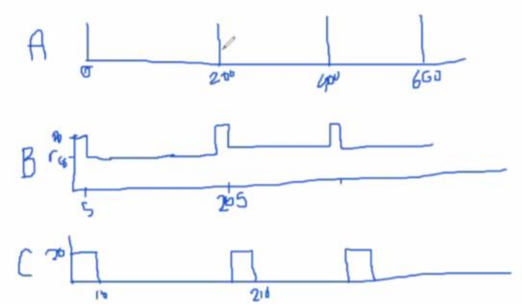
Neuron A sends excitatory inputs to neurons B and C (B and C are not directly connected).

- Neuron A fires regularly (a perfect pacemaker) at 5 spikes/sec
- Neuron B is a Poisson process with a rate of 60 spikes/s increasing to 80 spikes/s for a period of 5 ms following an incoming spike from neuron A.
- Neuron C is a Poisson process with a rate of 0 spikes/s increasing to 20 spikes/s for a period of 10ms following an incoming spike from neuron A.

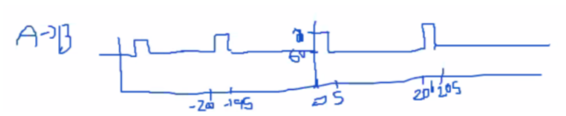
Draw the cross-correlation of A-B, A-C, B-C and C-B (in the range ±500ms).

Solution:

Illustration of the three neurons-



Cross correlation from A to B-



Similarly, cross correlation from A to C will look almost the same, but with highest of 0 and 20 and for periods of 10 ms.