5 3.25

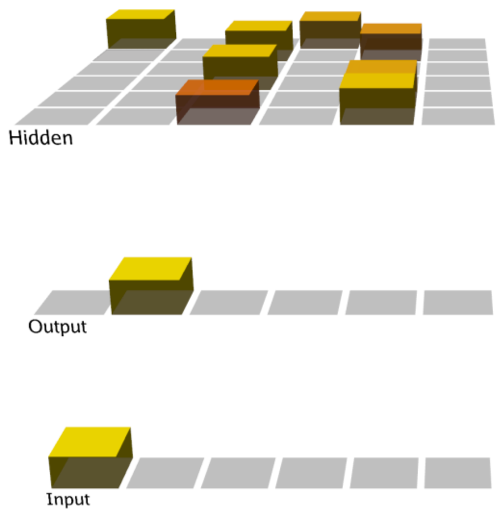
Computational Cognitive Neuroscience – Midterm Project

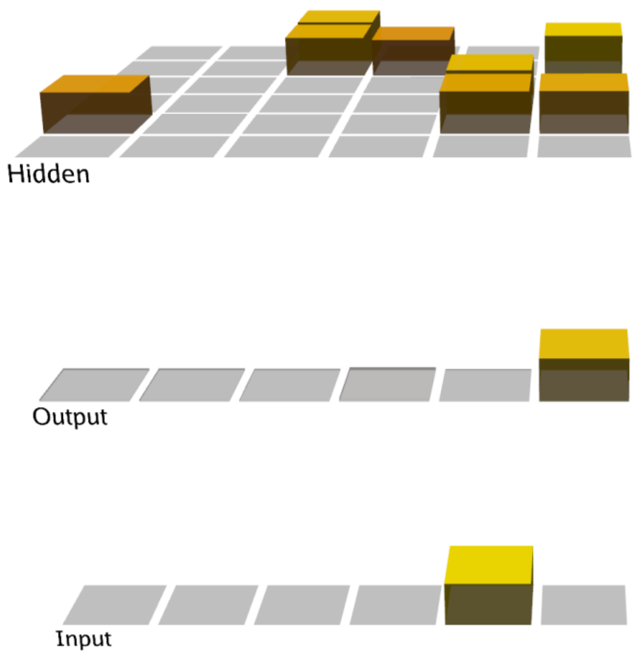
Zhuo Wang ScM BME Brown ID# 140641091

Question 1

It can learn to produce the correct output, given the input.

1





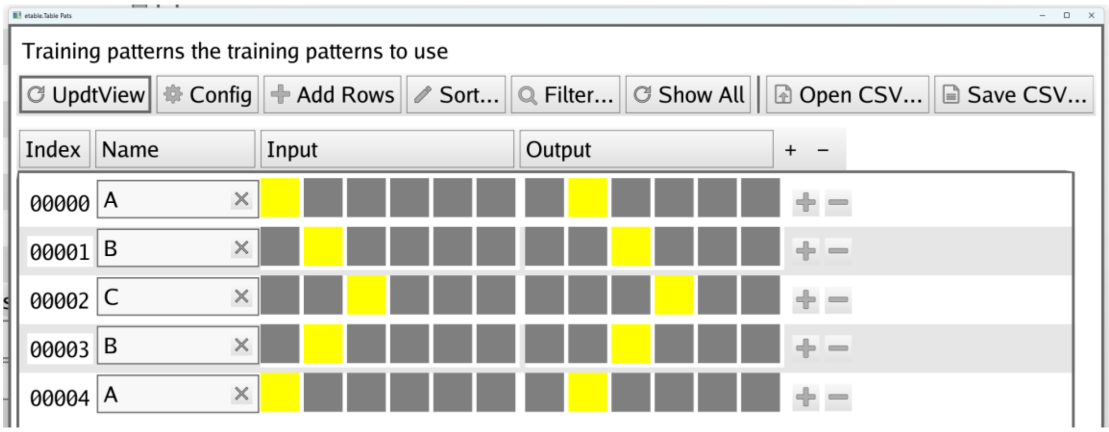
Question 2

I think it might be difficult to train successfully. Since the network would have to first
1 learn, how to translate the input sequence (ABCBA) to the matching output sequence.
There may be some error along the way. Keeping context across a lengthy series of steps is not a natural fit for the SRN. Every step of the network's feedforward

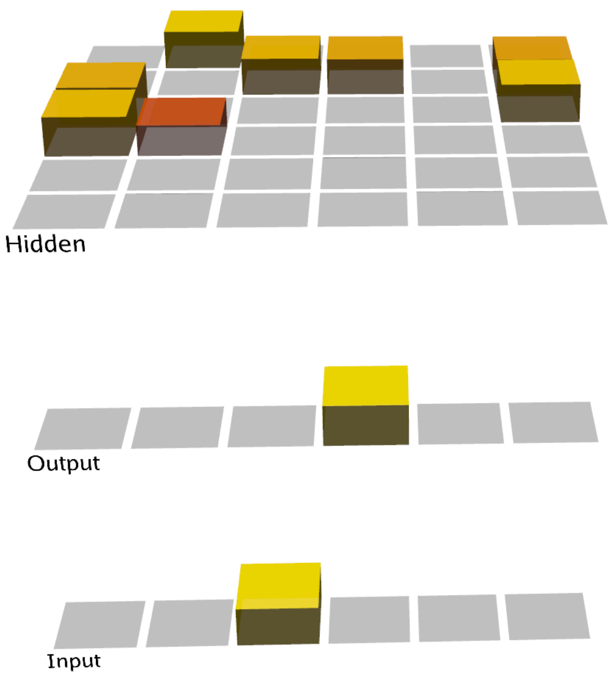
connections updates the activations according to the input received so far and the activations that came before it.

As the sequence forward from C to B and finally back to A, the hidden layer activations that were initially associated with A are likely to have shifted due to the recurrent activations. As a result of the recurrent dynamics overwriting the context in the hidden layer, the output generated for A can differ from the original intended output. The network has lost track of the original context, which could result in mistakes in the output.

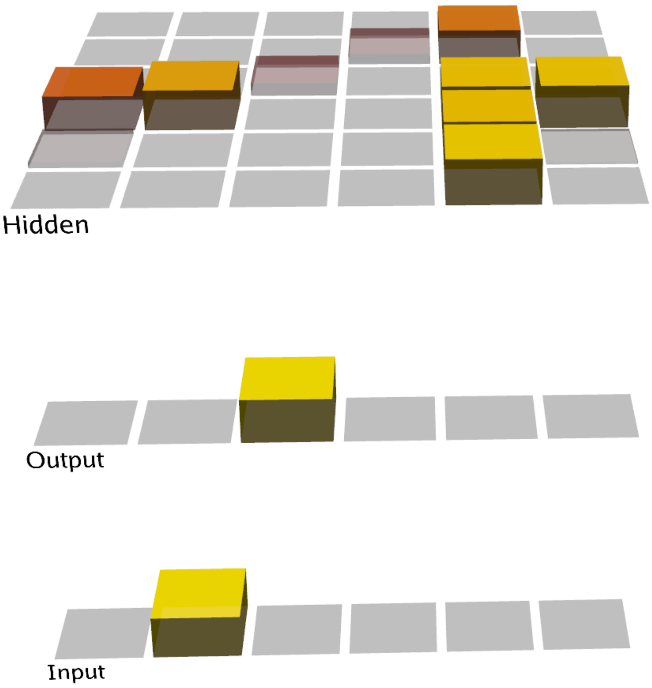
But my experiment turned out to be the opposite of what I expected, and the network successfully trained the correct output.



The sequence



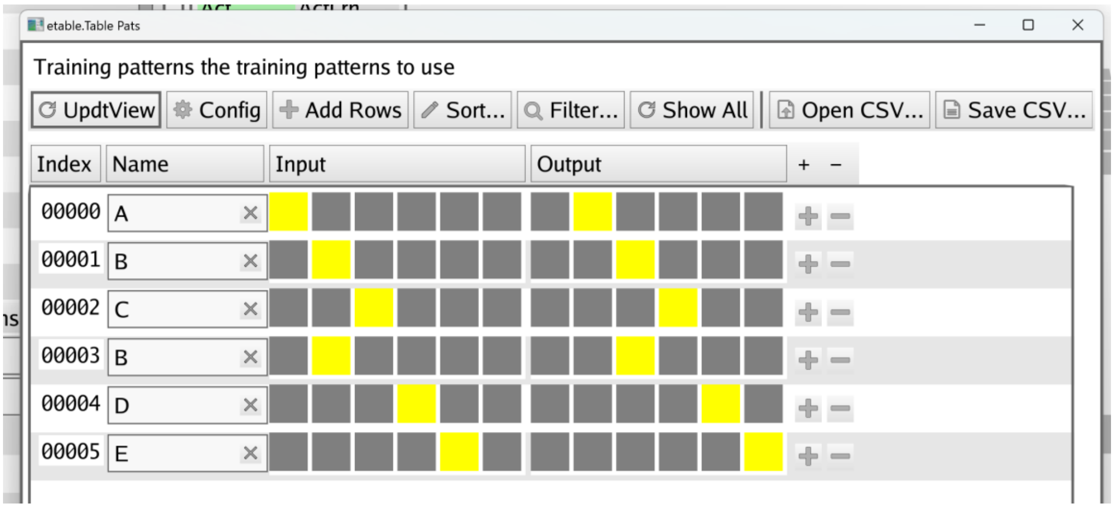
Test C



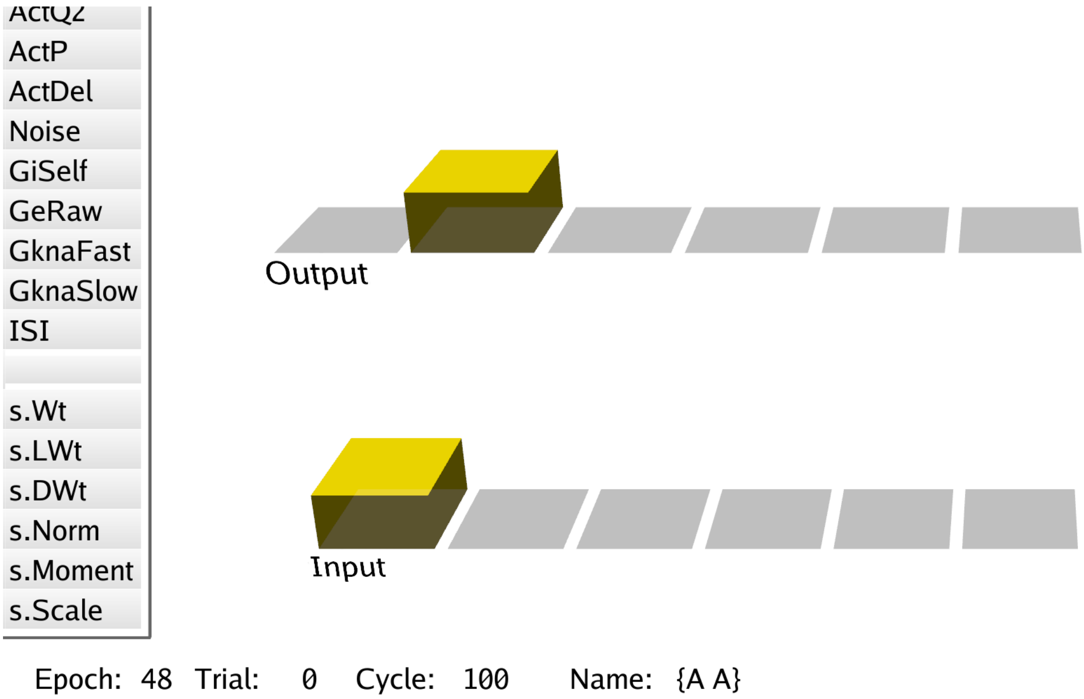
Test B

Question 3

From the experimental results, it can’t learn the sequence. Especially for BC and BD. This is due to SRN’s predominance of short-term memory. Although they may have
1 trouble remembering and retaining context over longer sequences, they are more equipped for capturing immediate dependencies in sequences. The network might not be able to maintain the required context in the cases of BC and BD, where information from previous inputs is essential for forecasting subsequent outputs.

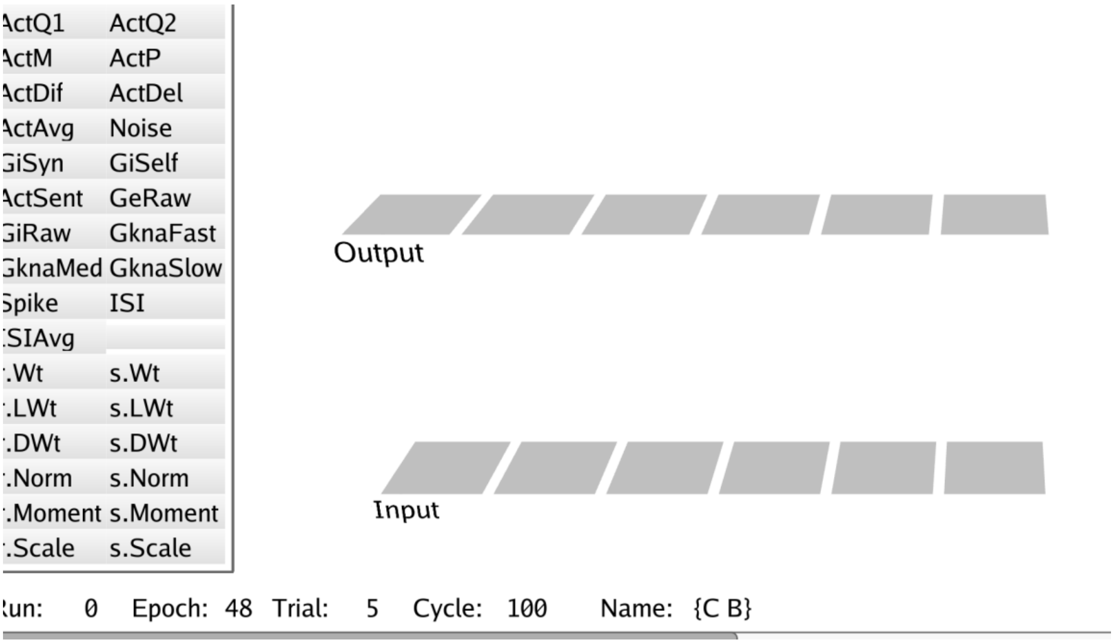


The sequence



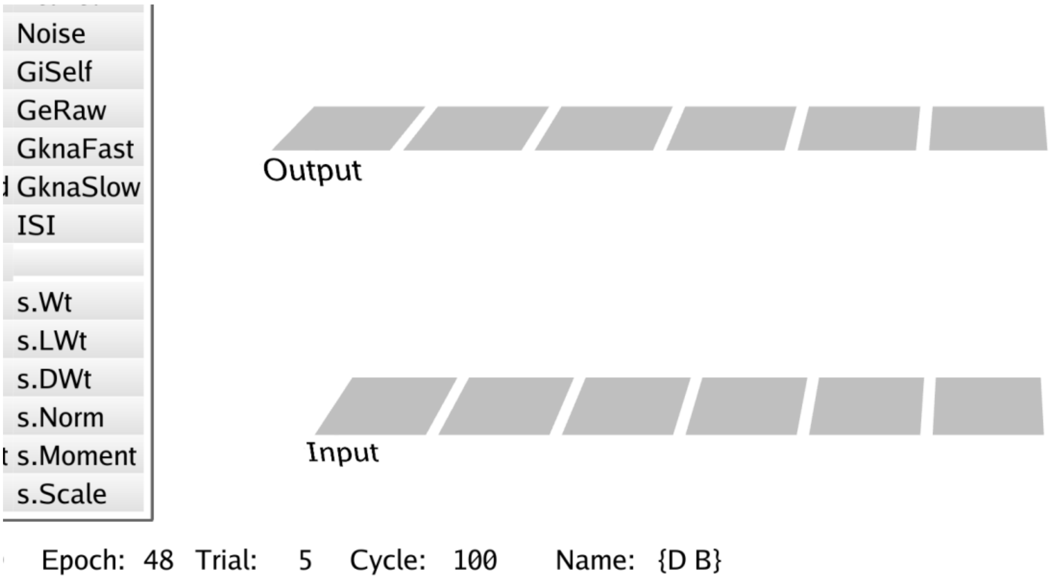
ActQ2
ActP
ActDel
Noise
GiSelf
GeRaw
GknaFast
GknaSlow
ISI
s.Wt
s.LWt
s.DWt
s.Norm
s.Moment
s.Scale
Output
Input
Epoch: 48 Trial: 0 Cycle: 100 Name: {A A}

A

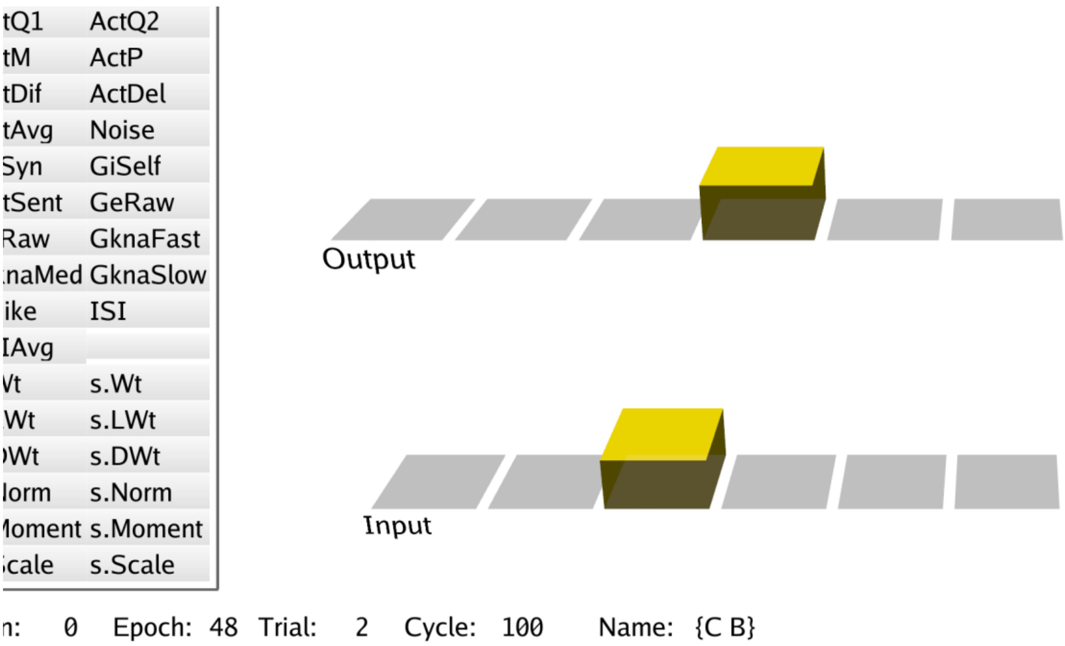


ActQ1 ActQ2
ActM ActP
ActDif ActDel
ActAvg Noise
GiSyn GiSelf
ActSent GeRaw
GiRaw GknaFast
GknaMed GknaSlow
Spike ISI
ISIAvg
.Wt s.Wt
.LWt s.LWt
.DWt s.DWt
.Norm s.Norm
.Moment s.Moment
.Scale s.Scale
Output
Input
Run: 0 Epoch: 48 Trial: 5 Cycle: 100 Name: {C B}

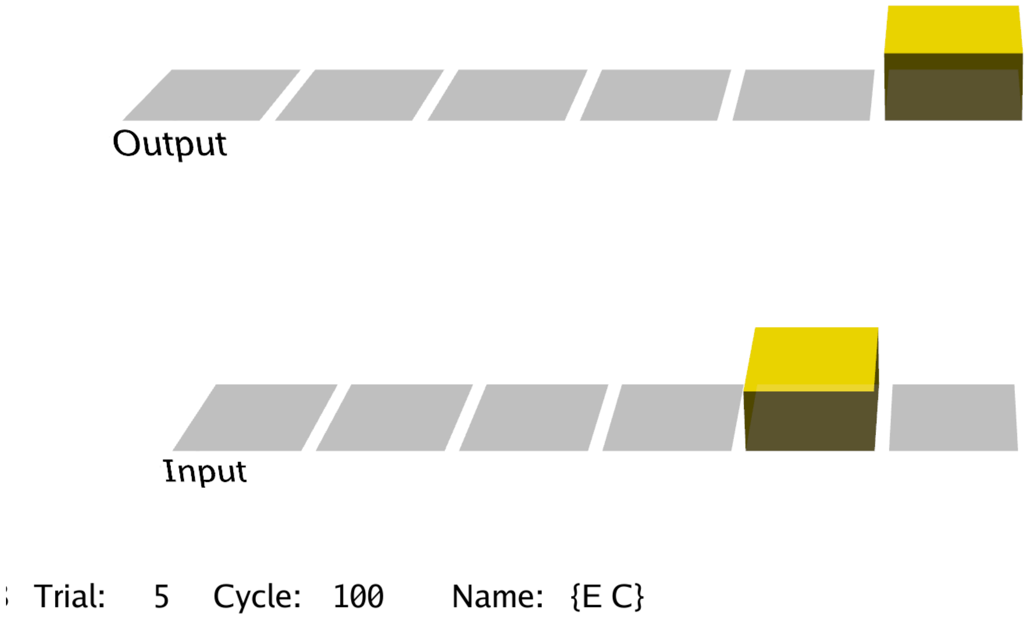
BC


Noise
GiSelf
GeRaw
GknaFast
GknaSlow
ISI
s.Wt
s.LWt
s.DWt
s.Norm
s.Moment
s.Scale
Output
Input
Epoch: 48 Trial: 5 Cycle: 100 Name: {D B}

BD


ActQ2
ActP
ActDel
Noise
GiSelf
GeRaw
GknaFast
GknaSlow
ISI
s.Wt
s.LWt
s.DWt
s.Norm
s.Moment
s.Scale
Output
Input
0 Epoch: 48 Trial: 2 Cycle: 100 Name: {C B}

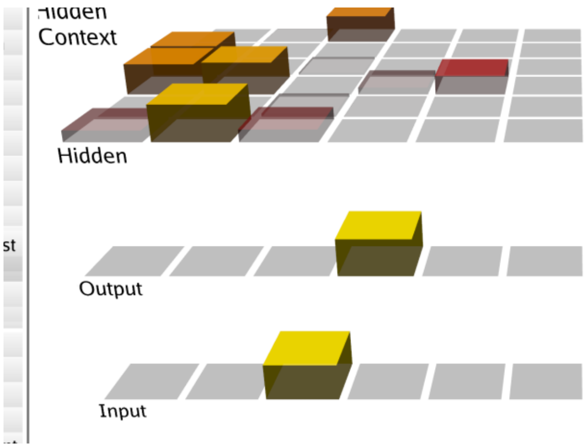
C



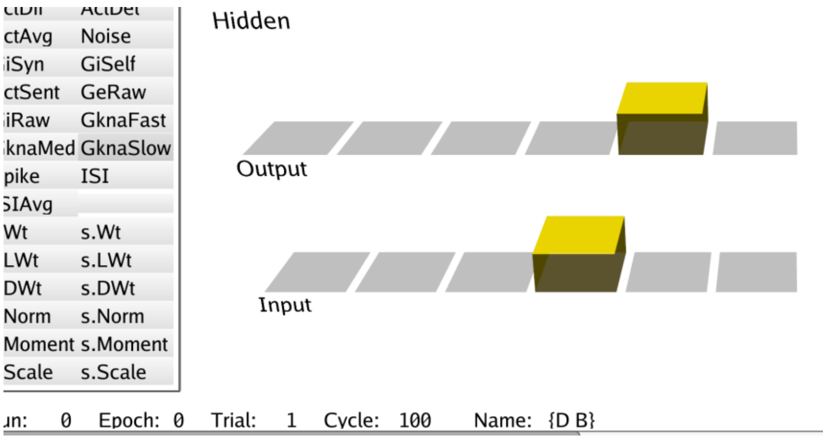
E

Question 4
It can learn the sequence.

The network can now preserve details about past trials, particularly the actions of the hidden layer, thanks to the addition of the "Hidden Context" layer. The activity of the
1 hidden layer is replicated to the context layer at the conclusion of each trial, therefore retaining a snapshot of the network's state. During the subsequent trial, predictions are made using the context provided by this memory. The network can use the context layer's stored data to make predictions in the current trial that take the impact of earlier inputs into consideration. This is particularly crucial in cases when the sequence of previous inputs determines the proper output, as in the case of the several "B" trials in "ABCBDE".



BC



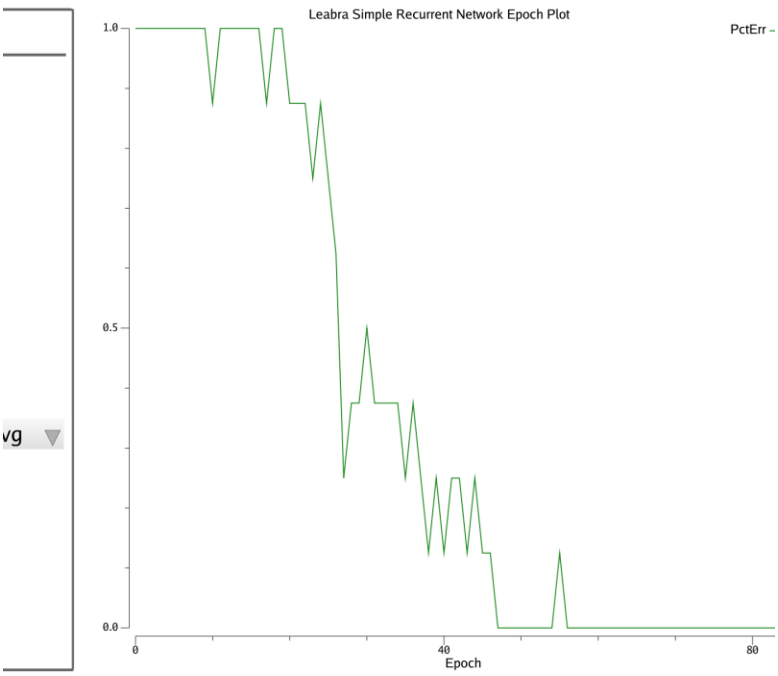
BD

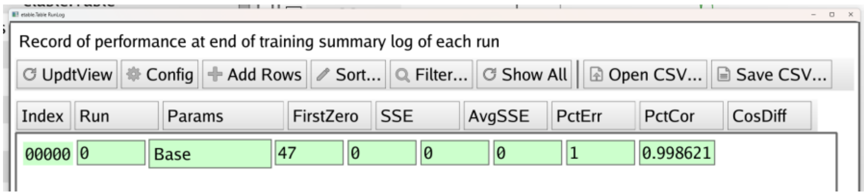
Question 5

The network can learn sequences.

The second-order dependency requires the network to remember not only the most recent trial but also the trial before that. In this instance, the network has to detect that the BC sequence was preceded by D in order to anticipate the right output when given E. The network's capacity to pick up such a sequence shows how well the training procedure and context layer capture higher-order dependencies. This implies that the network may now remember information from past trials and utilize it to predict future events based on longer-lasting patterns found in the input sequence.

47 epochs. (Screenshot below)





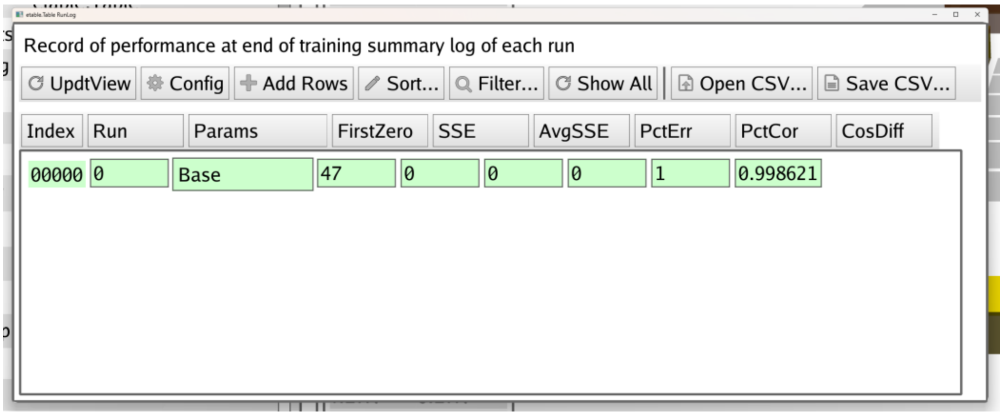
Record of performance at end of training summary log of each run

UpdtView | Config | Add Rows | Sort... | Filter... | Show All | Open CSV... | Save CSV...

Index	Run	Params	FirstZero	SSE	AvgSSE	PctErr	PctCor	CosDiff
00000	0	Base	47	0	0	0	1	0.998621

Question 6

According to the result, could complete the training under FmHid=1, FmHid=0 to FmHid=0.5, FmHid=0.5. The shortest training time is 29 epochs in the case of FmHid=0.7, FmHid=0.3.



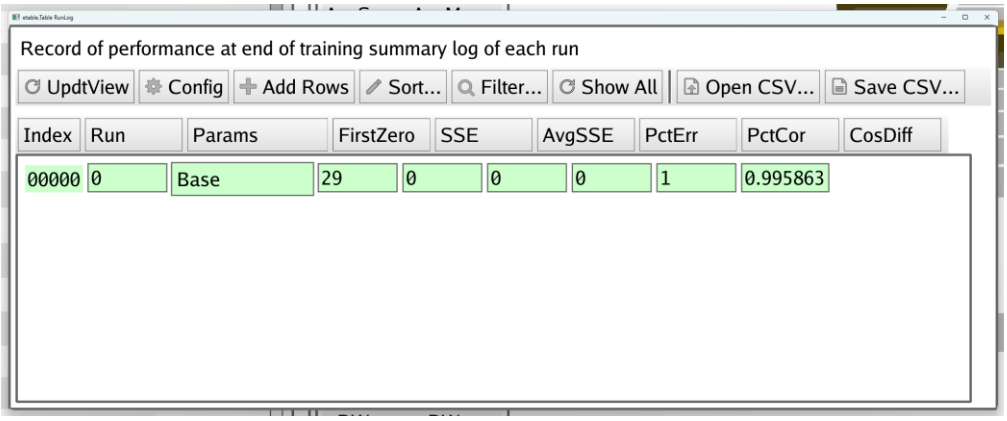
Record of performance at end of training summary log of each run

UpdtView | Config | Add Rows | Sort... | Filter... | Show All | Open CSV... | Save CSV...

Index	Run	Params	FirstZero	SSE	AvgSSE	PctErr	PctCor	CosDiff
00000	0	Base	47	0	0	0	1	0.998621

1

FmHid=1, FmHid=0

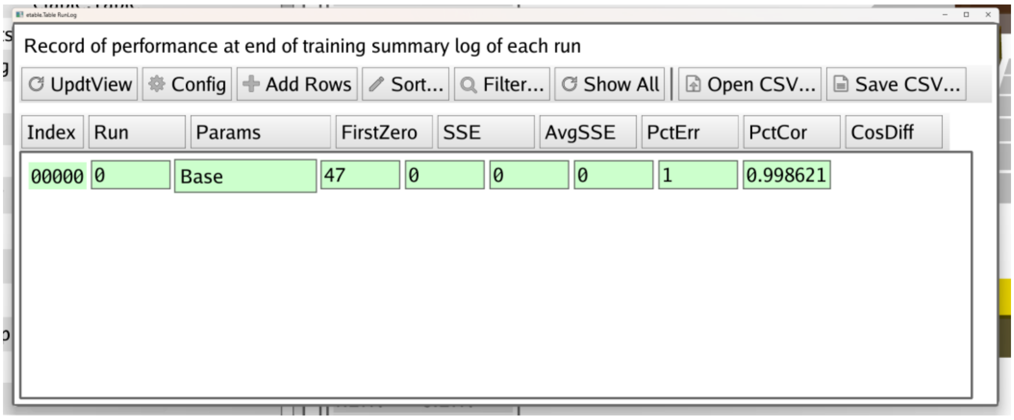


Record of performance at end of training summary log of each run

UpdtView | Config | Add Rows | Sort... | Filter... | Show All | Open CSV... | Save CSV...

Index	Run	Params	FirstZero	SSE	AvgSSE	PctErr	PctCor	CosDiff
00000	0	Base	29	0	0	0	1	0.995863

FmHid=0.7, FmHid=0.3


Record of performance at end of training summary log of each run

UpdtView Config Add Rows Sort... Filter... Show All Open CSV... Save CSV...

Index	Run	Params	FirstZero	SSE	AvgSSE	PctErr	PctCor	CosDiff
00000	0	Base	47	0	0	0	1	0.998621

FmHid=0.5, FmHid=0.5

Question 7
The way that data from earlier trials is incorporated into the present context is controlled by the parameters FmHid and FmPrv in the context layer of an SRN.

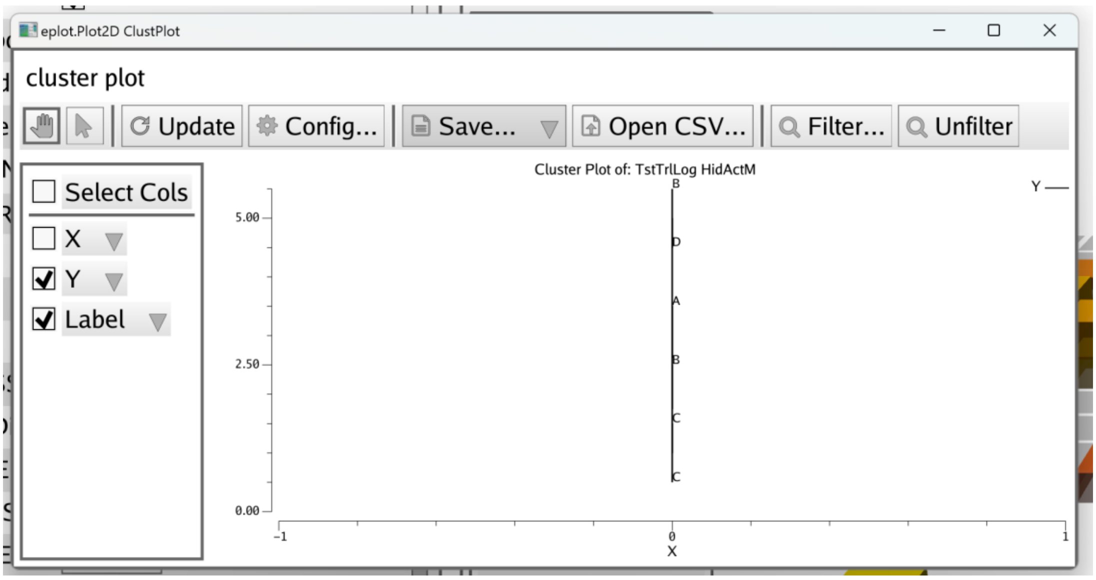
The context layer depends more on the data from the current hidden layer when FmHid is higher. This indicates that the network gives capturing current dependencies within the sequence more weight. Therefore, a higher FmHid value is appropriate when the network needs to focus on short-term dependencies and be more responsive
0.75 to the latest inputs.

The context layer depends more on its own prior context activity when FmPrv is higher. A higher FmPrv value is appropriate when the network needs to capture long-term dependencies from multiple trials. It helps the network remember and respond to patterns that go beyond the latest input.

Question 8
The patterns “ABC” and “DB” are clustered together. The reason they are clustered together is that they share a strong similarity in their representations in the hidden layer. In the context of the sequence, “ABC” serves as a contextual cue or input that predicts the subsequent pattern “DB”. The network learns to associate the “ABC” pattern with the following “DB” pattern. This association creates a similarity in their representations within the hidden layer because the network needs to remember the “ABC” context to produce the appropriate response “DB”. The “DB” is the most similar.

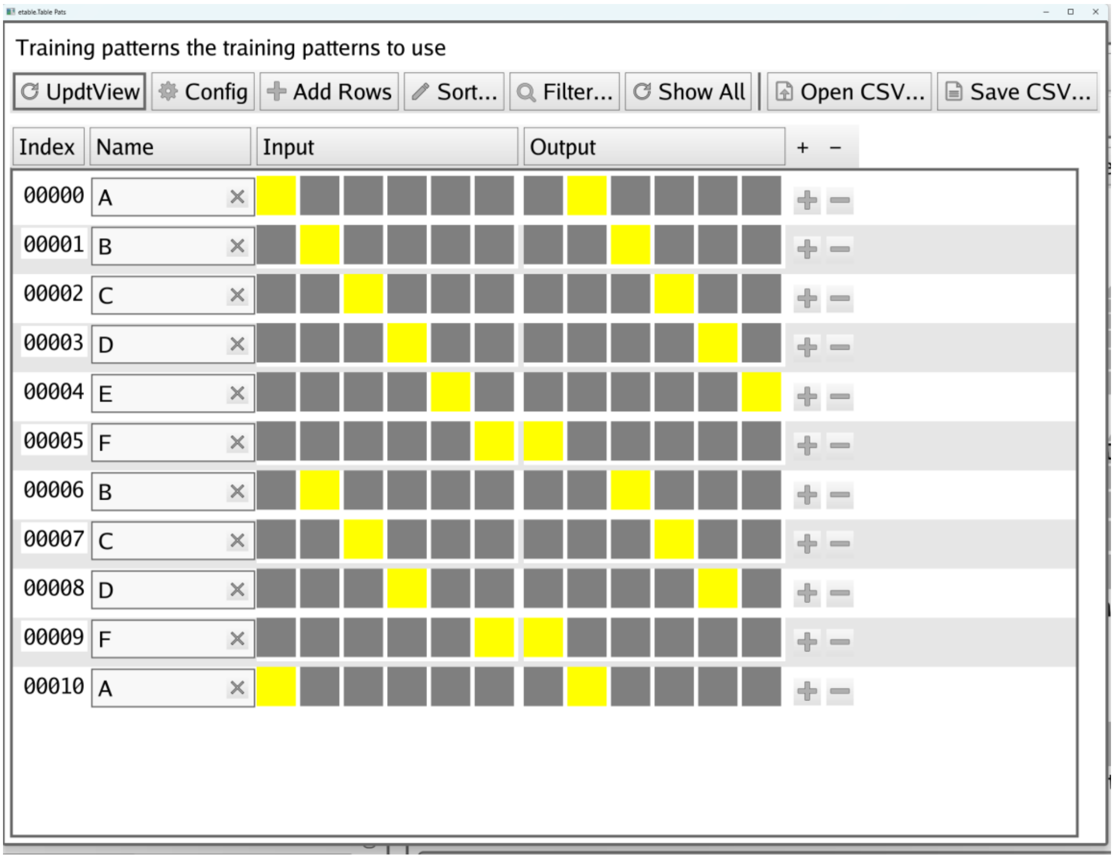
we should not expect to see a cluster plot like this. are you sure that you followed the appropriate instr

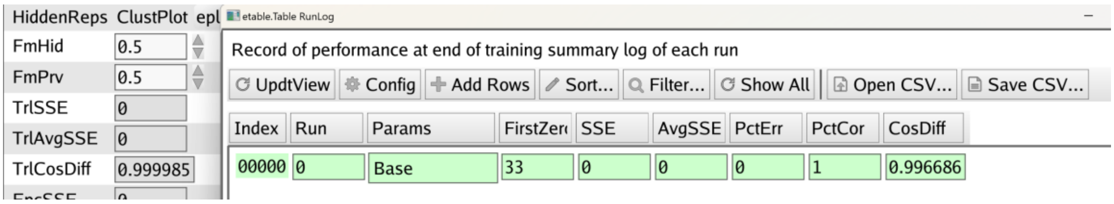
0



Question 9

I successfully tried fourth order dependency. I set the FmHid and FmPrv to 0.5. These values mean that the context layer combines information equally from the current hidden layer and the previous context layer's activity, allowing for a balanced integration of information from the past. Balancing these two parameters helps the
1 network maintain an appropriate level of context information, making it capable of learning and predicting patterns with a fourth order dependency.





Question 10
SRN can capture sequential dependencies and dynamics in data. My major is
0.5 biomedical engineering, and I think the causes of many chronic diseases are
cumulative, and we may be able to use this dependency to predict diseases. The risk
of lung cancer in smokers can be predicted by dependency on their age of smoking.