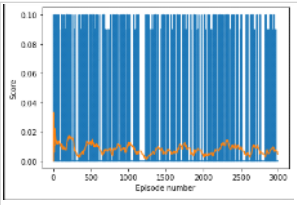
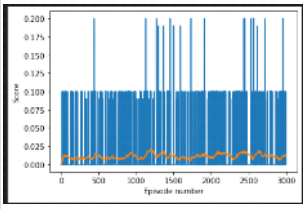
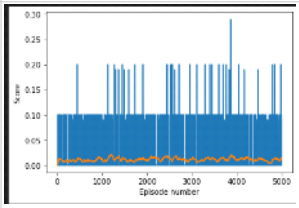
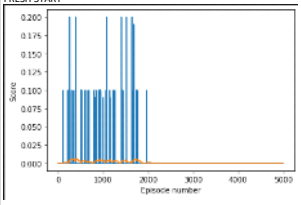
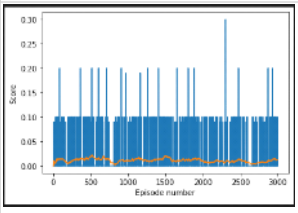
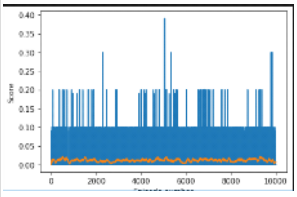
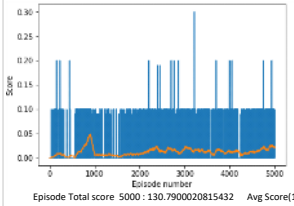
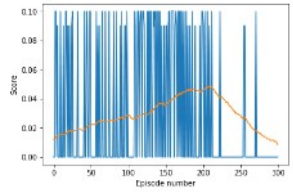
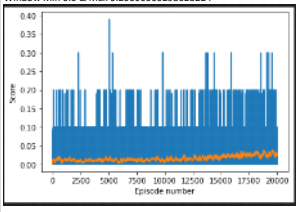
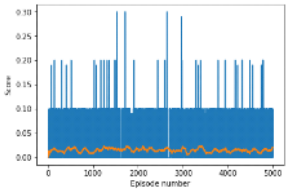
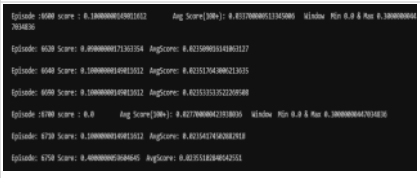
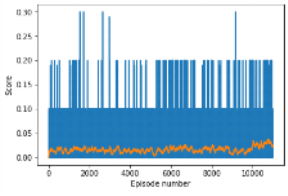
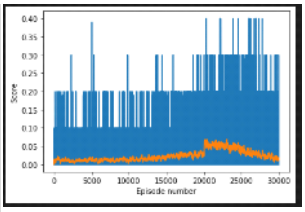
History

Thursday, 20 August 2020 5:45 PM

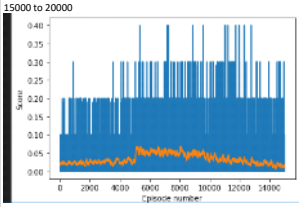
1.	#HP :Seed 1,Gamma :0.997 , TAU:0.0013, LR_Act :1.3e-05 , LR_Critic 1e-05 , Mu 0.0, Theta 0.17, Sigma 0.99, ExploreFactor 0.7, IsTargetHardcopyTrue #RL1.5 weight		
2.	HP :Seed 1,Gamma :0.997 , TAU:0.0013, LR_Act :1.3e-07 00000013 , LR_Critic 1e-07 , Mu 0.0, Theta 0.17, Sigma 0.99, ExploreFactor 0.7, IsTargetHardcopyTrue @500 avg 100+ **0.012** **@100 avg 100+ 0.0118** **RL1.5 Weight**	 Loaded weights from max 1.5	
3.	HP :Seed 1,Gamma :0.997 , TAU:0.0013, LR_Act :1.3e-07 (0.00000013), LR_Critic 1e-07 ,(0.0000001) Mu 0.0, Theta 0.17, Sigma 0.001, ExploreFactor 0.03, IsTargetHardcopyTrue @300 100+ 0.005 avg 0.002 Score - 0.6 @2000 Total score: 4.9 @4000 Total score :4.94 **FRESH START**	FRESH START  Relaoded with trained weigths (1.5) BUT NO SIGN OF LEARNIN so aborted the execution	
4.	HP :Seed 1,Gamma :0.997 , TAU:0.0013, LR_Act :1.3e-07 , LR_Critic 1e-07 , Mu 0.0, Theta 0.17, Sigma 0.79, ExploreFactor 0.9, IsTargetHardcopyTrue #R1.5' @300 = TS 3.03 avg 0.012 @400 max 0.2 avg 0.013 @1000 TS 11.72 100+ 0.009 Max @2000 TS 24 100+ 0.009 **@2400 max : 0.3 TS 27**	 Good sign - so continou for next 7000 	#Udacity Workspace #HP :Seed 1,Gamma :0.997999 , TAU:0.00013, LR_Act :2e-05 , LR_Critic 3e-05 , #Mu 0.0, Theta 0.77, Sigma 0.9, ExploreFactor 0.89, IsTargetHardcopyTrue  Episode Total score 5000 : 130.7900020815432 Avg Score(100+): 0.013800000213086605 Window Min 0.0 & Max 0.20000000298023224 zoomPlot(episodes_score,avg_score,300,700,1000) 

		Episode Total score 6300 : 221.94000348448753
		Avg Score(100+): 0.024700000379234553
		Window Min 0.0 & Max 0.20000000298023224
		 
		#Tune Hyper parameters - after 10K run
		#gut - steady learning but learning rate is slow so changing mainly Lr params. and exploit knowledge sinceconsisent avg getting 0.2 for every 100+
		setattr(colabAI,'gamma',0.998)
		setattr(colabAI,'tau',0.0013)
		setattr(colabAI,'explorfactor',0.6)
		setattr(colabAI,'lrActor',0.0001)
		setattr(colabAI,'lrCritic',0.0001)
			

	DYNAMIC TUNE
	1. Memory added only when Positve rewards after getting avg score 0.02
	2. Fine tune hyper parameters
	3. Clean up Old memories
	4. Changed np.max(agent_rewards) from sum To see the effects?
		
		Episode :9500 score : 0.0 Avg Score(100+): 0.021000000312924386 Window Min 0.0 & Max 0.20000000298023224 Episode: 9560 Score: 0.10000000149011612 AvgScore: 0.023516915112239824 Episode: 9600 Score: 0.10000000149011612 AvgScore: 0.023511824686822 76 Episode :9600 score : 0.10000000149011612 Avg Score(100+): 0.01990000003002584 Window Min 0.0 & Max 0.30000000447034836 Episode :9700 score : 0.0 Avg Score(100+): 0.013800000213086605 Window Min 0.0 & Max 0.30000000447034836 Episode: 9710 Score: 0.20000000298023224 AvgScore: 0.023477953912966764 Episode :9800 score : 0.0 Avg Score(100+): 0.016900000255554916 Window Min 0.0 & Max 0.30000000447034836 Episode :9900 score : 0.0 Avg Score(100+): 0.013000000193715095 Window Min 0.0 & Max 0.20000000298023224 Episode:
		
		9950 Score: 0.10000000149011612 AvgScore: 0.023409349275737353
		Episode :10000 score : 0.0 Avg Score(100+): 0.013900000210851431
		Window Min 0.0 & Max 0.30000000447034836
		
		REPLAY MEMORY - Refreshed
		 
			Episode: 2940 Score: 0.10000000149011612 AvgScore: 0.016553016714717667
			Episode: 2970 Score: 0.10000000149011612 AvgScore: 0.016561531710863765
			Episode: 2980 Score: 0.09000000171363354 AvgScore: 0.016563752538376983
			Episode :3000 score : 0.0 Avg Score(100+): 0.021700000334531 07 Window Min 0.0 & Max 0.20000000298023224
			 
			CLEARED OLD MEMORY
			
			#Tune Hyper parameters - after 10K run
			#gut - steady learning but learning rate is slow so changing mainly Lr params. and exploit knowledge sinceconsisent avg getting 0.3 a
			setattr(colabAI,'gamma',0.978)
			setattr(colabAI,'tau',0.0013)
			setattr(colabAI,'explorfactor',0.78)



15000 to 20000

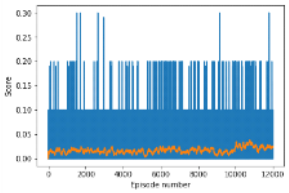


MEMORY REFRESHED

```
#Tune Hyper parameters - after 10K run
#steady learning but learning rate is slow so changing mainly Lr params. and  exploit knowledge sinceconsisent avg getting 0
.2 for every 100+
setattr(colabAI,'gamma',0.9989)
setattr(colabAI,'tau',0.0013)
setattr(colabAI,'explorfactor',0.57)
setattr(colabAI,'lrActor',0.00013)
setattr(colabAI,'lrCritic',0.0001)
#Original : HP :Seed 1,Gamma :0.997 , TAU:0.0013, LR_Act :1.3e-07 , LR_Critic 1e-07 , Mu 0.0, Theta 0.17, Sigma 0.79, Explor
eFactor 0.9, IsTargetHardcopyTrue
```

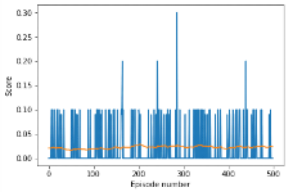
setattr(colabAI,'tau',0.0013)
setattr(colabAI,'explorfactor',0.78)
setattr(colabAI,'lrActor',0.001)
setattr(colabAI,'lrCritic',0.001)

Ran for 1K



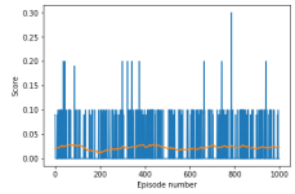
Episode: 920 Score: 0.10000000149011612 AvgScore: 0.01702181234996416
Episode: 940 Score: 0.20000000298023224 AvgScore: 0.01704355135850196
Episode :1000 score : 0.0 Avg Score(100+): 0.023600000366568567 Window Min 0.0 & Max 0.20000000298023224

From <https://viewka0rezltofl.udacity-student-workspaces.com/notebooks/Untitled%20Folder%202/Tennis.ipynb>



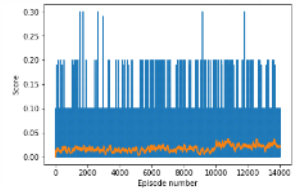
Above - last 500

```
zoomPlot(episodes_score,avg_score,1000,11000,12000)
```



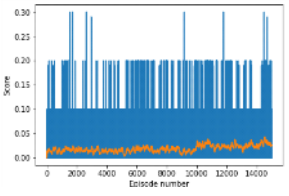
#Tune Hyper parameters - after 10K run
#gut - steady learning but learning rate is slow so changing mainly Lr params. and exploit knowledge sinceconsisent avg getting 0.3 a
setattr(colabAI,'gamma',0.978)
setattr(colabAI,'tau',0.0013)
setattr(colabAI,'explorfactor',0.78)
setattr(colabAI,'lrActor',0.0001)
setattr(colabAI,'lrCritic',0.0001)
setattr(colabAI,'theta',0.24)
setattr(colabAI,'sigma',0.89)

Cud see max per eepisode reduced to 0.2
So increase Gamma
#Tune Hyper parameters - after 10K run
#Gamma gets lower - consistency max to 0.2 so increase gamma back to 0.99
#3rd change
setattr(colabAI,'gamma',0.9998)
setattr(colabAI,'tau',0.0013)
setattr(colabAI,'explorfactor',0.78)
setattr(colabAI,'lrActor',0.0001)
setattr(colabAI,'lrCritic',0.0001)
setattr(colabAI,'theta',0.24)
setattr(colabAI,'sigma',0.89)



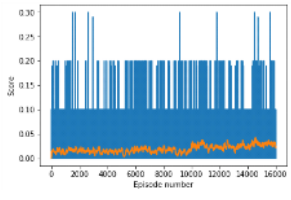
Changed - Memory experience to have np.max(rewards) instead of np.sum(rewards) -

#Tune Hyper parameters - after 10K run
#Gamma gets lower - consistency max to 0.2 so increase gamma back to 0.99
#4th time tuning back tau 0.013 and explore to 0.978
setattr(colabAI,'gamma',0.9998)
setattr(colabAI,'tau',0.013)
setattr(colabAI,'explorfactor',0.978)
setattr(colabAI,'lrActor',0.0001)
setattr(colabAI,'lrCritic',0.0001)
setattr(colabAI,'theta',0.24)
setattr(colabAI,'sigma',0.89)



Np.max

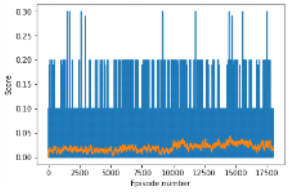
#Tune Hyper parameters - after 10K run
#Gamma gets lower - consistency max to 0.2 so increase gamma back to 0.99
#5th time tuning on lr and explore
setattr(colabAI,'gamma',0.9998)
setattr(colabAI,'tau',0.013)
setattr(colabAI,'explorfactor',0.78)
setattr(colabAI,'lrActor',0.0000001)
setattr(colabAI,'lrCritic',0.0000013)
setattr(colabAI,'theta',0.24)
setattr(colabAI,'sigma',0.89)

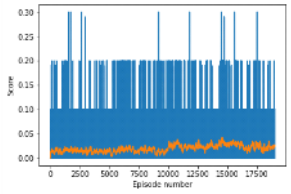


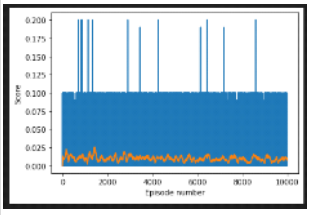
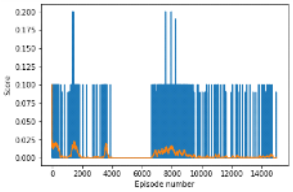
Np.sum

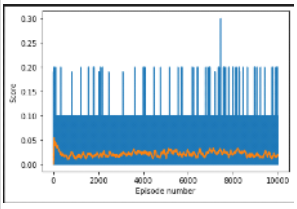
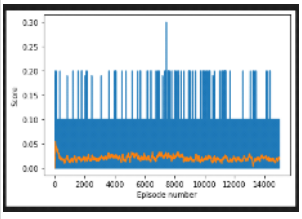
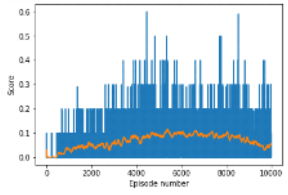
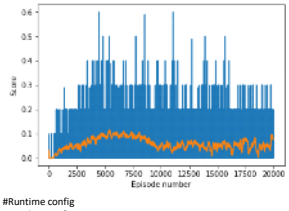
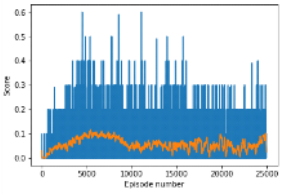
```
Reward add to memory if its greater than 0.1 instead of 0.3
#Tune Hyper parameters - after 10K run
#3rd time change
# Lr rate and explore facotor
```

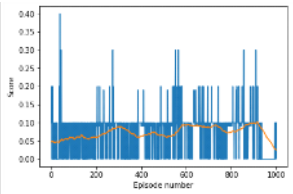
#Tune Hyper parameters - after 10K run
#Gamma gets lower - consistency max to 0.2 so increase gamma back to 0.99
#6th time tuning on lr to too low and explore to 0.6
setattr(colabAI,'gamma',0.9998)

		-Restore back since not much learning only exploring
		
		`setattr(colabAI,'gamma',0.99897)`
		`setattr(colabAI,'tau',0.0013)`
		`setattr(colabAI,'explorfactor',0.69)`
		`setattr(colabAI,'lrActor',0.0000013)`
		`setattr(colabAI,'lrCritic',0.0000001)`
		`setattr(colabAI,'sigma',0.82)`
		`setattr(colabAI,'theta',0.24)`
		`#Original : HP :Seed 1,Gamma :0.997 , TAU:0.0013, LR_Act :1.3e-07 , LR_Critic 1e-07 , Mu 0.0, Theta 0.17, Sigma 0.79, Explor eFactor 0.9, IsTargetHardcopyTrue`
			setattr(colabAI,'tau',0.013)
			setattr(colabAI,'explorfactor',0.6)
			setattr(colabAI,'lrActor',0.000000001)
			setattr(colabAI,'lrCritic',0.000000013)
			setattr(colabAI,'theta',0.24)
			setattr(colabAI,'sigma',0.89)
			 

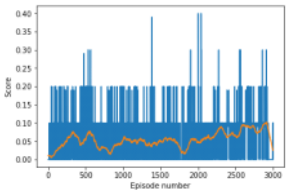
			#Back to original setting #7th time setattr(colabAI,'gamma',0.9998) setattr(colabAI,'tau',0.013) setattr(colabAI,'explorfactor',0.978) setattr(colabAI,'lrActor',0.0001) setattr(colabAI,'lrCritic',0.0001) setattr(colabAI,'theta',0.24) setattr(colabAI,'sigma',0.89) 

	WINDOWS	HP :Seed 1,Gamma :0.997 , TAU:0.0013, LR_Act :1.3e-07 , LR_Critic 1e-07 , Mu 0.0, Theta 0.17, Sigma 0.79, ExploreFactor 0.9, IsTargetHardcopyTrue
		 
			#HP :Seed 1,Gamma :0.987999 , TAU:0.00013, LR_Act :2e-06 , LR_Critic 3e-06 ,
			#Mu 0.0, Theta 0.77, Sigma 0.9, ExploreFactor 0.79, IsTargetHardcopyTrue
			showGraph(episodes_score,avg_score)
			 

MAX .6
 	
EFFECT OF MEMORY CHANGES
	Loaded the network weights From Udacity wos trained	`#HP :Seed 1,Gamma :0.997999 , TAU:0.0013, LR_Act :0.00013 , LR_Critic 0.0001 , Mu 0.0, Theta 0.17, Sigma 0.9, ExploreFactor 0.79, IsTargetHardcopyFalse`
		
		Episode :100 score : 0.0 Avg Score(100+): 0.04180000068619847 Window Min 0.0 & Max 0.20000000298023224
		 
		
		NO MAJOR IMPROVEMENT so try Memory refresh
		
		`#common Replay Memory for both Agents`
		`#Refresh Memory`
		`#Delete all old memories since mostly 0.00 or negative experience stored in last 20K`
		`#Try with current network but reduced Buffer size so very old memories will be erased`
		`#replayMemory = ReplayBuffer(action_size, BUFFER_SIZE, BATCH_SIZE, SEED)`
		`#Reduced memory size 5M to 0.2M so only recent expereinces retained- old memory archived not deleted`
		`replayMemory_bkup5M = replayMemory`
		`setattr(colabAI.commonMemory,"memory",deque(maxlen=100000))`
		`setattr(colabAI.commonMemory,"batch_size",128)`
		
		And ran for 5K
		 
		
		No improvement - change HP
		`#replayMemory_bkup5M = replayMemory`
		`setattr(colabAI.commonMemory,"memory",deque(maxlen=100000))`
		`setattr(colabAI.commonMemory,"batch_size",256)`
		`#Training Tennis Agents`
		`#HP :Seed 1,Gamma :0.99997 , TAU:0.0017, LR_Act :0.0013 , LR_Critic 0.0013 , Mu 0.0, Theta 0.17, Sigma 0.79, ExploreFactor 0 .69, IsTargetHardcopyTrue RE-RUN Hyper param tuned`
		
		RUNNNING
			`#HP :Seed 1, SEEDC=5`
			Gamma :0.997999 ,
			TAU:0.0013,
			LR_Act :0.0002 , LR_Critic 0.0003 ,
			Mu 0.0, Theta 0.77, Sigma 0.9, ExploreFactor 0.79,
			IsTargetHardcopyTrue
			
			Episode: 4470 Score: 0.60000000089406967 AvgScore: 0.04740715956927946
			Episode: 4490 Score: 0.10000000149011612 AvgScore: 0.04768151521334271
			Episode :4500 score : 0.0 Avg Score(100+): 0.09330000141635537 Window Min 0.0 & Max 0.60000000089406967
			
			FIRST TIME Avg Score +0.1
			pisode :6000 score : 0.10000000149011612 Avg Score(100+): 0.10390000162646175 Window Min 0.0 & Max 0.30000000447034836
			
			Episode :6600 score : 0.10000000149011612 Avg Score(100+): 0.10490000167861581 Window Min 0.0 & Max 0.30000000447034836
			
			Episode :6900 score : 0.10000000149011612 Avg Score(100+): 0.09580000150948763 Window Min 0.0 & Max 0.4000000059604645
			
			Episode :7800 score : 0.20000000298023224 Avg Score(100+): 0.08640000130981207 Window Min 0.0 & Max 0.5000000074505806
			
			Episode :8600 score : 0.0 Avg Score(100+): 0.0884000013396144 Window Min 0.0 & Max 0.5000000091642141
			 
			
			Continues 10K
			 
			
			#Runtime config
			#1st change after 20K - 21 start
			#just explore factor reduced & theta 0.001
			setattr(colabAI,'gamma',0.997999)
			setattr(colabAI,'tau',0.0013)
			setattr(colabAI,'explorfactor',0.69)#reduced
			setattr(colabAI,'lrActor',0.0002)
			setattr(colabAI,'lrCritic',0.0003)
			setattr(colabAI,'sigma',0.9)
			setattr(colabAI,'theta',0.76)#reduced
			 
			
			ZOOM 24k to 25k



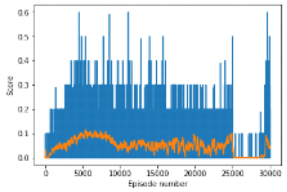

```
zoomPlot(episodes_score,avg_score,3000,22000,25000)
```



Same kind of learning or little down
Why cant refresh the moemry
- Old memories with more negative rcords may recall-
- If we empty past memories and with new learning will get add more positive rewards greater than 0. or 0.2

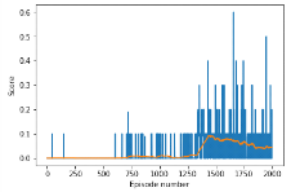
SO, Changed to see the effects
#common Replay Memory for both Agents
#Refresh Memory
#Delete all old memories since mostly 0.00 or negative experience stored in last 20K
#Try with current network but reduced Buffer size so very old memories will be erased
#replayMemory = ReplayBuffer(action_size, BUFFER_SIZE, BATCH_SIZE, SEED)
#Reduced memory size 5M to 0.2M so only recent expereinces retained - old memory archived not deleted
replayMemory_bkup5M = replayMemory
setattr(colabAI.commonMemory,"memory",deque(maxlen=200000))
setattr(colabAI.commonMemory,"batch_size",256)

#MEMORY

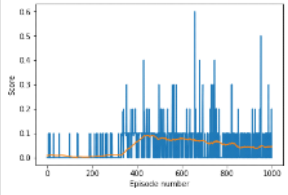


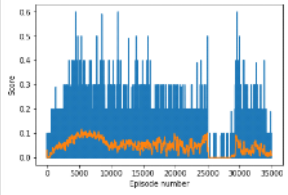
#ZOOM Last 2000

```
#after memory refershed
zoomPlot(episodes_score,avg_score,2000,28000,30000)
```

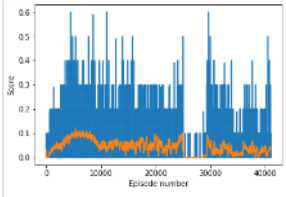


Last 1000 below

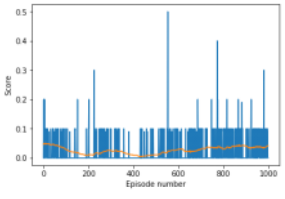




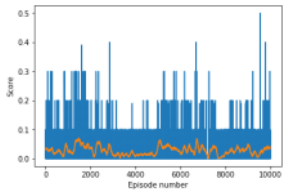
IMPROVEMENT

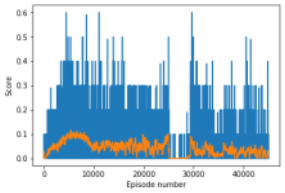



```
zoomPlot(episodes_score,avg_score,1000,40000,41000)
```



```
zoomPlot(episodes_score,avg_score,10000,31000,41000)
```






```
print('Length of Episodes {} and Max Score so far:{} '.format(len(episodes_score),np.max(episodes_score)))
```

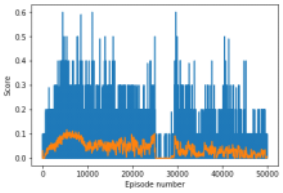
```
Length of Episodes 45000 and Max Score so far:0.6000000089406967
```

#just gamma reduced from .997999 to .987999
setattr(colabAI,'gamma',0.987999)#reduced
setattr(colabAI,'tau',0.0013)
setattr(colabAI,'explorfactor',0.89)
setattr(colabAI,'lrActor',0.00021) #0.01
setattr(colabAI,'lrCritic',0.00031)#0.01
setattr(colabAI,'sigma',0.9)
setattr(colabAI,'theta',0.79)

#Training Tennis Agents
##HP :Seed 1,Gamma :0.987999 , TAU:0.0013, LR_Act :0.00021 , LR_Critic 0.00031 ,
#Mu 0.0, Theta 0.77, Sigma 0.9, ExploreFactor 0.79, IsTargetHardcopyTrue

#45K completed w- Continue for next 5k

```
Length of Episodes 50000 and Max Score so far:0.600000008940696
```

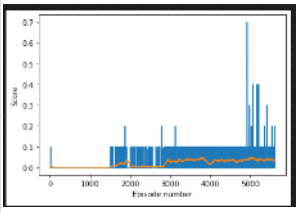


#backup recent memories
replayMemory_bkup30to50k = colabAI.commonMemory
#reset the memory size
setattr(colabAI.commonMemory,"memory",deque(maxlen=5000000))
#restore old memories
setattr(colabAI,"commonMemory",replayMemory_bkup5M)
setattr(colabAI.commonMemory,"batch_size",256)

WINDOWS

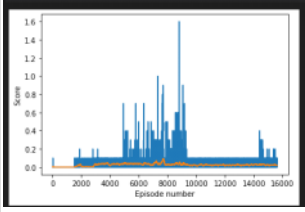
```
{
    "SEEDC": 2,
    "SEED" : 1,
    "BUFFER_SIZE": 1000000,
    "BATCH_SIZE": 256,
    "GAMMA": 0.9387999,
    "TAU": 0.0011,
    "LR_CRITIC": 0.00019,
    "LR_ACTOR": 0.00019,
    "MU":0,
    "THETA": 0.26,
    "SIGMA" : 0.36,
    "EXPLORE" :0.03



}
"SEEDC": 2,
    "SEED" : 1,
    "BUFFER_SIZE": 1000000,
    "BATCH_SIZE": 256,
    "GAMMA": 0.9387999,
    "TAU": 0.0011,
    "LR_CRITIC": 0.00019,
    "LR_ACTOR": 0.00019,
    "MU":0,
    "THETA": 0.26,
    "SIGMA" : 0.36,
    "EXPLORE" :0.03

#Load trained model weights where got 0.5 in 1800 episodes itself
actornnk_path ="./Maxpoint9/checkpoint_actor_point_base-copy4.pth"
criticnn_path ="./Maxpoint9/checkpoint_critic_point_base-copy4.pth"
act_tar_path = "./Maxpoint9/checkpoint_actor_target_base-copy4.pth"
critc_tar_path="./Maxpoint9/checkpoint_critic_target_base-copy4.pth"
colabAI.load_trained_model(actornnk_path,criticnn_path,act_tar_path,critc_tar_path)


#tune 2 after 5k episodes
for agent in colabAI.multiagent:
    setattr(agent,"gamma",0.9986)
    setattr(agent,"explorfactor",0.003)
    setattr(agent,"lrActor",0.000055)
    setattr(agent,"lrCritic",0.00003)
    setattr(agent,"tau",0.0013)
    setattr(agent,"sigma",0.369)
    setattr(agent,"theta",0.39)
```

Episode: 4310 Score: 0.10000000149011612 AvgScore (100+): 0.03570000058040023 CurrentWindowMax: 0.10000000149011612 Episode: 4320 Score: 0.0 AvgScore (100+): 0.03820000059902668 CurrentWindowMax: 0.7000000104308128

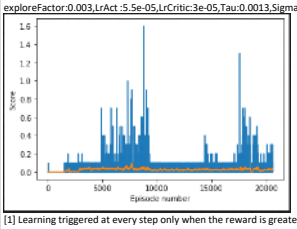


Episode: 720 Score: 0.0 AvgScore (100+): 0.029900000449270 01 CurrentWindowMax: 0.30000000447034836 Episode: 730 Score: 0.7000000104308128 AvgScore (100+): 0.037900000568479 3 CurrentWindowMax: 0.7000000104308128 Episode: 740 Score: 0.0 AvgScore (100+): 0.041900000628083946 CurrentWindowMax: 0.7000000104308128

```
#HP :Seed 1,Gamma :0.9986 , TAU:0.0013, LR_Act :0.000055 , LR_Critic 0.00003 , Mu 0.0, Theta 0.39, Sigma 0.369
ExploreFactor 0.003, IsTargetHardcopyFalse
showGraph(episodes_score,avg_score)
```



Gamma:0.9986 , exploreFactor:0.003,LrAct :5.5e-05,LrCritic:3e-05,Tau:0.0013,Sigma:0.369,Theta:0.39 Gamma:0.9986 , exploreFactor:0.003,LrAct :5.5e-05,LrCritic:3e-05,Tau:0.0013,Sigma:0.369,Theta:0.39



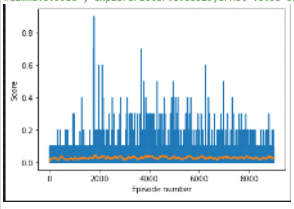
[1] Learning triggered at every step only when the reward is greater than 0.19 .
After 20k episodes when getting consistent 0.3 to 0.7 to 1.0
Changed the learning every step to see the effect -11.45 am 29th aug

[2]

```
if(np.max(agent_scores)>=0.39):
#changed from 0.19 to 0.39 after 15k epsidoes -got max 1.7
 replayHighMemoryRewards.add(states,actions,rewards,next_states,dones)
```

[3]

```
#Gamma:0.9986 , exploreFactor:0.00329,LrAct :5.5e-06,LrCritic:3e-06,Tau:0.0013,Sigma:0.169,Theta:0.039
```



Gamma:0.99896 , exploreFactor:0.00327,LrAct :7.5e-08,LrCritic:2e-07,Tau:0.00013,Sigma:0.279,Theta:0.19

```
#Load trained model weights where got 0.5 in 1800 episodes itself
actornnk_path ="./Max1_7/checkpoint_actor_point_base1p7.pth"
criticnn_path ="./Max1_7/checkpoint_critic_point_base1p7.pth"
act_tar_path = "./Max1_7/checkpoint_actor_target_base1p7.pth"
critc_tar_path="./Max1_7/checkpoint_critic_target_base1p7.pth"
colabAI.load_trained_model(actornnk_path,criticnn_path,act_tar_path,critc_tar_path)
```

5K - Learn only when reward >0.19

Then 6k to 10 - every step

Episode: 660 Score: 0.0 AvgScore (100+): 0.027800000421702862 CurrentWindowMax: 0.6000000089406967
Episode: 670 Score: 0.10000000149011612 AvgScore (100+): 0.040800000615417956 CurrentWindowMax: 1.0000000149011612

Episode: 860 Score: 0.0 AvgScore (100+): 0.027000000402331352 CurrentWindowMax: 0.4000000059604645 Episode: 870 Score: 0.0 AvgScore (100+): 0.049000000730156895 CurrentWindowMax: 1.8000000268220901

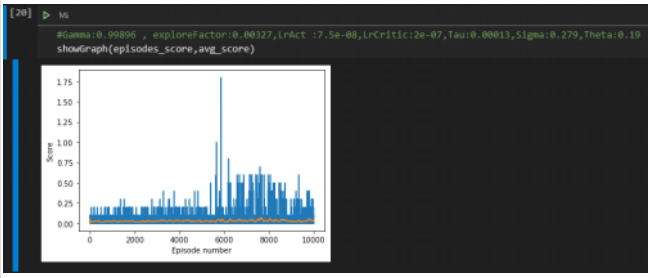
```
Episode :800 score : 0.0         Avg Score(100+): 0.03690000055357814  WinMax 0.4000000059604645

Episode: 810 Score: 0.0  AvgScore (100+): 0.03390000058887466 CurrentWindowMax: 0.4000000059604645

Episode: 820 Score: 0.0  AvgScore (100+): 0.03290000049397349 CurrentWindowMax: 0.4000000059604645

Episode: 830 Score: 0.0  AvgScore (100+): 0.027900000419467687 CurrentWindowMax: 0.4000000059604645

Episode: 840 Score: 0.0  AvgScore (100+): 0.029000000432133674 CurrentWindowMax: 0.4000000059604645

Episode: 850 Score: 0.0  AvgScore (100+): 0.030000000447034835 CurrentWindowMax: 0.4000000059604645

Episode: 860 Score: 0.0  AvgScore (100+): 0.027000000402331352 CurrentWindowMax: 0.4000000059604645

Episode: 870 Score: 0.0  AvgScore (100+): 0.049000000730156895 CurrentWindowMax: 1.8000000268220901

[ ]  ▷ M↓
   #Gamma:0.99896 , exploreFactor:0.00327,LrAct :7.5e-08,LrCritic:2e-07,Tau:0.00013,Sigma:0.279,Theta:0.19
   showGraph(episodes_score,avg_score)
```

```
Episode :800 score : 0.0         Avg Score(100+): 0.03690000055357814  WinMax 0.4000000059604645

Episode: 810 Score: 0.0  AvgScore (100+): 0.03390000058887466 CurrentWindowMax: 0.4000000059604645

Episode: 820 Score: 0.0  AvgScore (100+): 0.03290000049397349 CurrentWindowMax: 0.4000000059604645

Episode: 830 Score: 0.0  AvgScore (100+): 0.027900000419467687 CurrentWindowMax: 0.4000000059604645

Episode: 840 Score: 0.0  AvgScore (100+): 0.029000000432133674 CurrentWindowMax: 0.4000000059604645

Episode: 850 Score: 0.0  AvgScore (100+): 0.030000000447034835 CurrentWindowMax: 0.4000000059604645

Episode: 860 Score: 0.0  AvgScore (100+): 0.027000000402331352 CurrentWindowMax: 0.4000000059604645

Episode: 870 Score: 0.0  AvgScore (100+): 0.049000000730156895 CurrentWindowMax: 1.8000000268220901

Episode: 880 Score: 0.10000000149011612  AvgScore (100+): 0.045000000067055225 CurrentWindowMax: 1.8000000268220901

[*]  ▷ M↓
   #Gamma:0.99896 , exploreFactor:0.00327,LrAct :7.5e-08,LrCritic:2e-07,Tau:0.00013,Sigma:0.279,Theta:0.19
   showGraph(episodes_score,avg_score)
```

Episode: 2170 Score: 0.10000000149011612 AvgScore (100+): 0.040000000059604645 CurrentWindowMax: 0.6000000089406967
Episode: 2180 Score: 0.0 AvgScore (100+): 0.041000000061094761 CurrentWindowMax: 0.6000000089406967
Episode: 2190 Score: 0.0 AvgScore (100+): 0.041000000061094761 CurrentWindowMax: 0.6000000089406967
Episode: 2200 Score: 0.4000000059604645 AvgScore (100+): 0.047000000070035457 CurrentWindowMax: 0.6000000089406967
Episode :2200 score : 0.4000000059604645 Avg Score(100+): 0.047000000070035457 WinMax 0.6000000089406967
Episode: 2210 Score: 0.0 AvgScore (100+): 0.043000000064074993 CurrentWindowMax: 0.6000000089406967

```
Episode: 2590 Score: 0.0  AvgScore (100+): 0.045000000670552225 CurrentWindowMax: 0.5000000074505806
Episode: 2600 Score: 0.0  AvgScore (100+): 0.046000000685453415 CurrentWindowMax: 0.7000000104308128
Episode :2600 score : 0.0        Avg Score(100+): 0.046000000685453415  WinMax 0.7000000104308128
Episode: 2610 Score: 0.0  AvgScore (100+): 0.060000000894069670 CurrentWindowMax: 0.7000000104308128
Episode: 2620 Score: 0.0  AvgScore (100+): 0.059000000879168510 CurrentWindowMax: 0.7000000104308128
Episode: 2630 Score: 0.10000000149011612  AvgScore (100+): 0.060000000894069670 CurrentWindowMax: 0.7000000104308128
Episode: 2640 Score: 0.0  AvgScore (100+): 0.058000000864267350 CurrentWindowMax: 0.7000000104308128
Episode: 2650 Score: 0.0  AvgScore (100+): 0.060000000894069670 CurrentWindowMax: 0.7000000104308128
Episode: 2660 Score: 0.0  AvgScore (100+): 0.060000000894069670 CurrentWindowMax: 0.7000000104308128
Episode: 2670 Score: 0.0  AvgScore (100+): 0.059000000879168510 CurrentWindowMax: 0.7000000104308128
Episode: 2680 Score: 0.6000000089406967  AvgScore (100+): 0.063000000938773160 CurrentWindowMax: 0.7000000104308128
```

```
Episode: 2970 Score: 0.0  AvgScore (100+): 0.026000000387430190 CurrentWindowMax: 0.4000000059604645
Episode: 2980 Score: 0.0  AvgScore (100+): 0.035000000521540640 CurrentWindowMax: 0.6000000089406967
Episode: 2990 Score: 0.5000000074505806  AvgScore (100+): 0.044000000655651090 CurrentWindowMax: 0.6000000089406967
Episode: 3000 Score: 0.0  AvgScore (100+): 0.050000000745058060 CurrentWindowMax: 0.6000000089406967
Episode :3000 score : 0.0        Avg Score(100+): 0.050000000745058060  WinMax 0.6000000089406967
Episode: 3010 Score: 0.0  AvgScore (100+): 0.051000000759959224 CurrentWindowMax: 0.6000000089406967
```



Refreshed memory length 4640

Re-run for next 5000