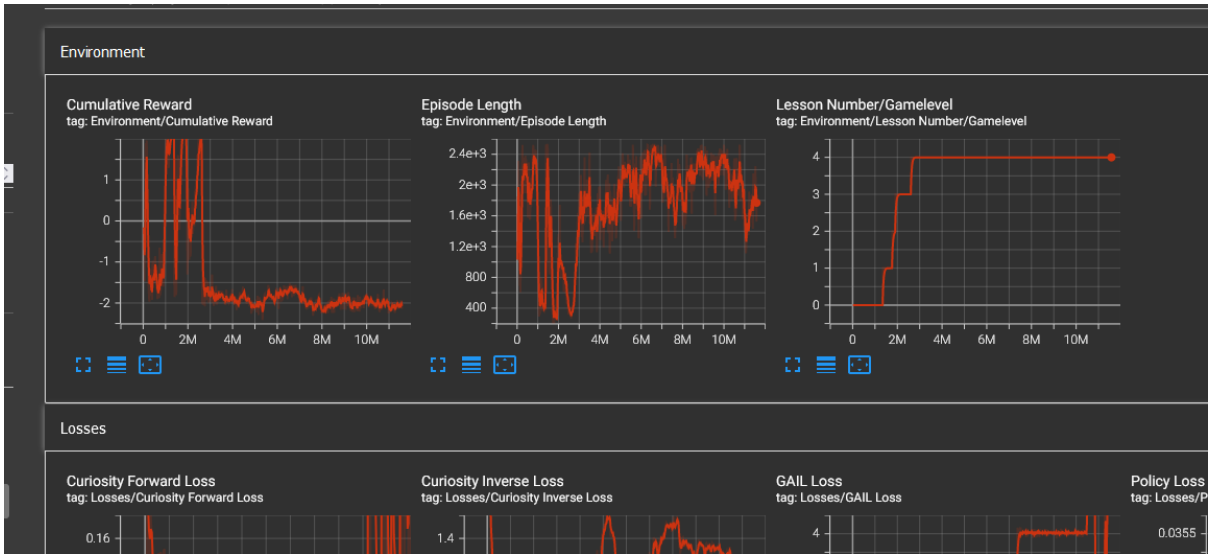
CyberMan ML progress

Run3 : Can get up to Level 5, but then cannot learn any Better behaviour at Level 5.

The Level 3 was too easy, Quick Transition.



Changes for Run 4

Hyper Parameter Changes

- Reduce Curiosity Strength from 0.05 down to 0.02
- Reduce GAIL Strength from 0.15 down to 0.1
- Increase Batch Size from 512 to 1024 (c.f. similar to Dalek ML)
- Increase Buffer Size from 5120 to 10240
- Reduce Learning Rate from 0.0003 to 0.00015

Change the Levels

- Level 3 (Was too easy) - Adjusted spawn Positions
- Level 5 (Seems too much of a challenge) - Adjusted Spawn Positions
- Level 7 is now the same as previous Level 5 Challenge
- All levels after 7 are Random.

Run4 : Notice that Player Ray Casts NOT Sensing the probe Collision Collider during game play. But is Sensed during Out of Play.

Seems to sense Probe OK during Hueristic Play. But adjust Start vertical Height of Ray to 1.5 on all.

But Noted that Daleks do not appear to have any Tags? So needed to Set Tags for All Daleks.

Seems to generally detect Probes, but NOT always, on some environments the Rays pass right through, not Sense, but then discovers/ sense on later the Probe manoeuvres ?

Run 4 Cannot get Past Level 2 !! - Rays are not hitting the Probe.

Run 5

Increase the size of the Probe Collider size

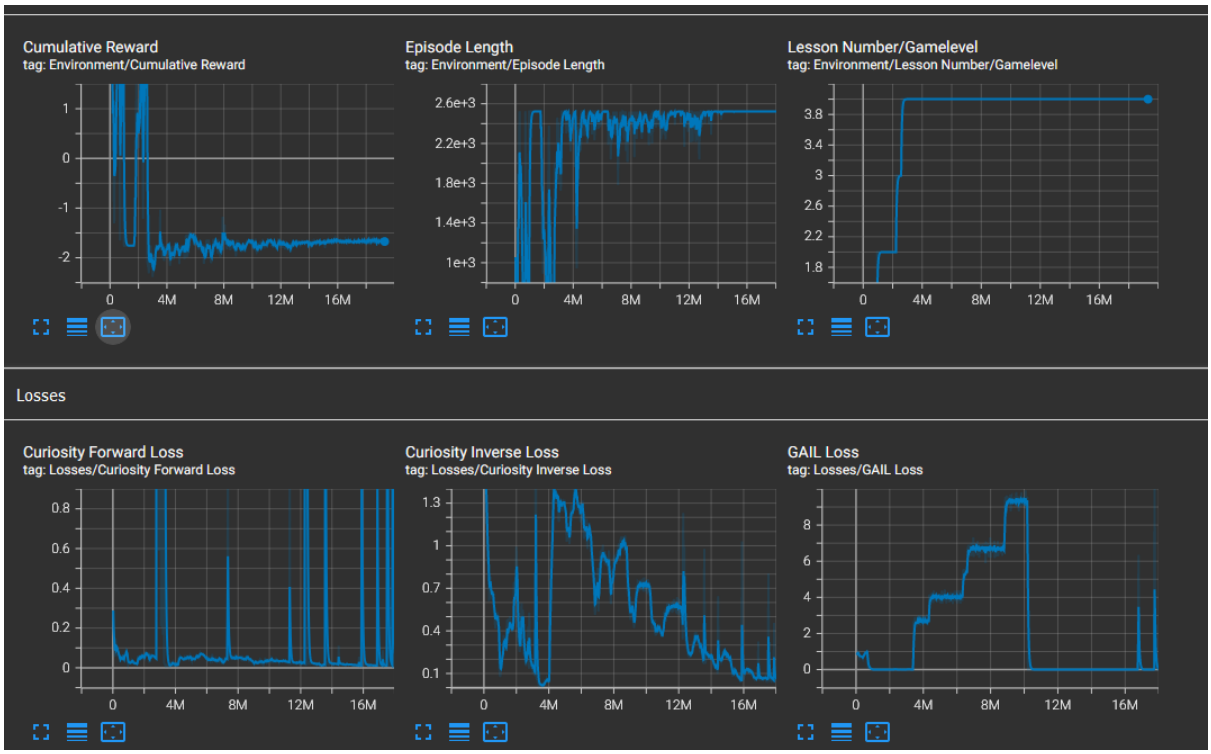
Revert Back some of the PPO Hyper parameters

- Increase GAIL Strength back to 0.15
- Revert Batch Size , and Reduce to 256
- Revert Buffer Size to 5120
- Increase Learning Rate to 0.0002
- Increase the Time_Horizon (from 128) to 512

**** Bug:** Note If the Episode aborts whilst holding the Reactor, the Collider will remain Disabled – hence the reason for Not Seeing the Probe.

So need to adjust the Probe code, to accept It Being Reset, Collider On, at Beginning of Each Episode.

Run 5 gets stuck on Level 5 – little progress after 10 Million Steps in same level.



Run 6 – Major Simplification

- Collapse the Actions into a single Branch – size 5, So Single Branch: Action [0] NOOP, [1], Fwd, [2] Left, [3] Right, [4] Pickup, [5] Place Down
- Reduce Observations remove the spatial Vector differences to Dalek, Goal, and Probe, so rely upon Ray Casts Only. Size/Shape of Observation => 3

- Remove GAIL learning . Rationale is that there are major GAIL losses in Level 5, but the training was not relevant there.

See if learn the basic levels without GAIL, and solely upon Ray Casts sensors?

Fails to learn any moves – Picks up but no Place downs in Goals

Maybe put GAIL Back in ? - as will struggle on that first lesson !

Run 7

** Bug – Player Branch Size should be 6 ! (not 5 – as per previous run) – There are 6 discrete Actions

Record a Demo2, and reinstate GAIL at Strength 0.2

Did not get beyond Level 2, after 5 Million Runs

Run 8

Disable GAIL

Did not get out of Level 1 ! - seems as though we still need GAIL

Somehow we need to recover Run 3 and Run 5 Performance

Run 9

- Reinstate 3 Branch Actions with Sizes: [2], [3], [3]
- *(But still leave Observations as mainly Ray casts)*
- Re perform Demos BasicTraning3
- Reinstate GAIL
- Reduce the Time_Horizon (from 512) to 256

Can't get beyond Level 3

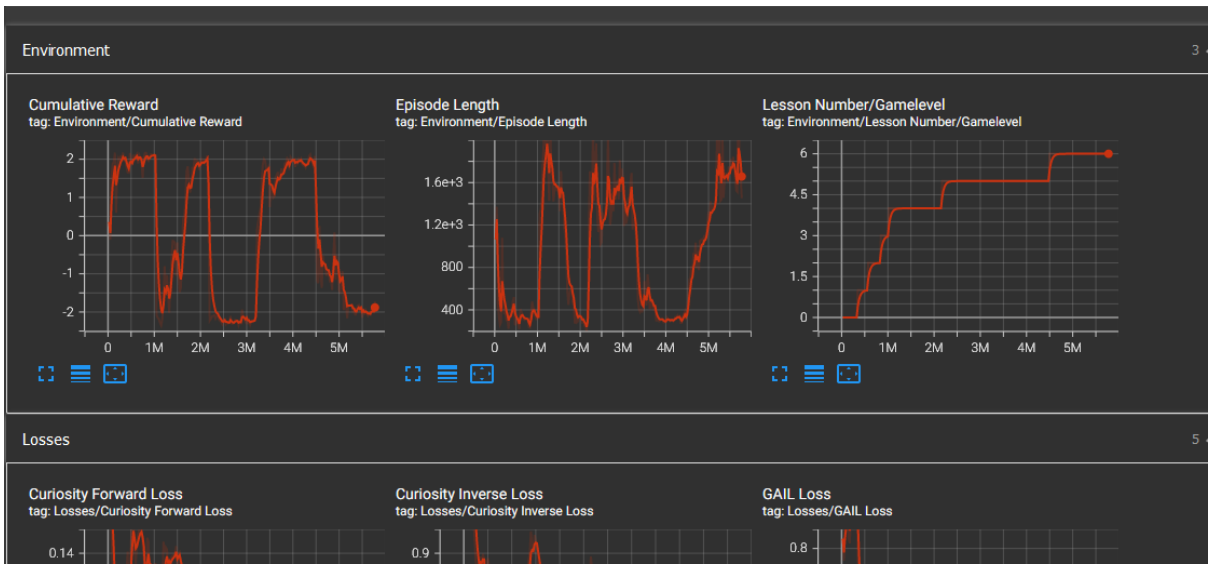
Run 10

Need to adjust Reward Assignments, Add Reward Immediate upon Condition, rather than at End of Animation, End Episode, so that recognise Action Credit quicker.

- Reduce batch size to 256
- Increase Model Network => 512 Hidden Nodes (from 256)
- The Inference Device, in Player can be set to GPU (as per Digby Game)

Cyberman Agent now gets to Level 5, in less than 1 Million Steps. Reward and performance drop off the cliff at level 5. No longer a difficult level, but the Initial Agent behaviour, circles around in main arena, allows Dalek to catch Agent. – Need some discovery of goal area. Does figure out Level 5, and transitions to level 6 at around 2.2 Million Steps.

Got to Level 7 at around 5 Million. However notice that some of the Rudium Reactors are being Lost, overboard, Not in the game !



Run 11,

Need to add some protection on Rudium Reactor not going overboard !

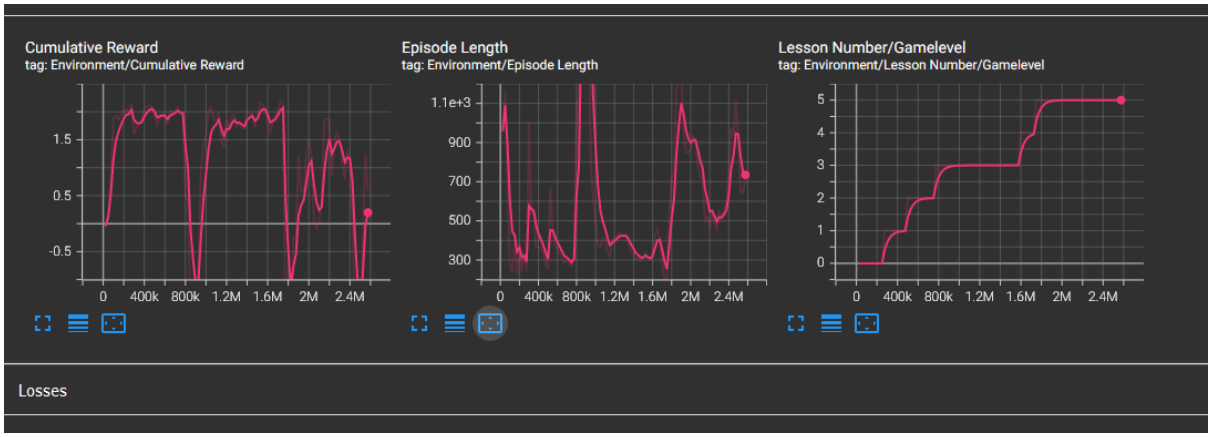
Clear Rudium Velocities at Start of Episode, so Can recover the Rudium Probe.

Run 11 Gets to Level 7 earlier, at around 3 Million Steps. But is Struggling at this level, still on Level 7 at 7 Million Steps

Still Stuck on Level 7, after 30 Million Steps ! Fails to Solve this Level

Run 12

- Change definition of Level 7
- Increase Beta from 0.001 to 0.005 (More Entropy)
- Increase curiosity strength from 0.02 to 0.04
- Programmed Levels 8,9,10. So now only Level 11 is Random Play. Level 10 similar to Previous Level 5/7 Difficult Level



Still getting stuck on Level 7, which implies that it is more a function of the Dalek Speed perhaps.

Check evaluation of ONXX file, at 4 Million Steps, pretty good at Level 4

This Brain works reasonably well at Level 1, but not so Well on Level 2, nor 3.

Backup to CyberManMLC

Run20

Revised Spawn environments, such that Uniform Random Spawn Positions, with Wider Extent as levels rise. So Extent = Level*[Uniform.Random(0:10) -5.0]/ 50.0 for Level 1..10

And LERP interpolate, between [(x1,z1), (x2,z2),(x3,z3),(x4,z4)] points, origin at (x3-x2)/2, (z3-z2)/2

Dalek spawns the North, Player to the South, with the Rubium probe starting in the East side and the Goal towards the West.

Halve the speed of Daleks, But Increase Initial Laser Range to 7.5

Seems to learn Levels 1 through to level 5 pretty quickly, in 700,000 steps, and to level 7 within 1 Million Steps. So the agent is finding it easier to progress through the smaller changes in Scenarios. With little drop off in any Rewards, between the Level Transitions. (I guess the Random placements, average out the new larger spawn displacements challenges)

Daleks seem to be going Backwards on level 3! - perhaps appears due to really low speed.

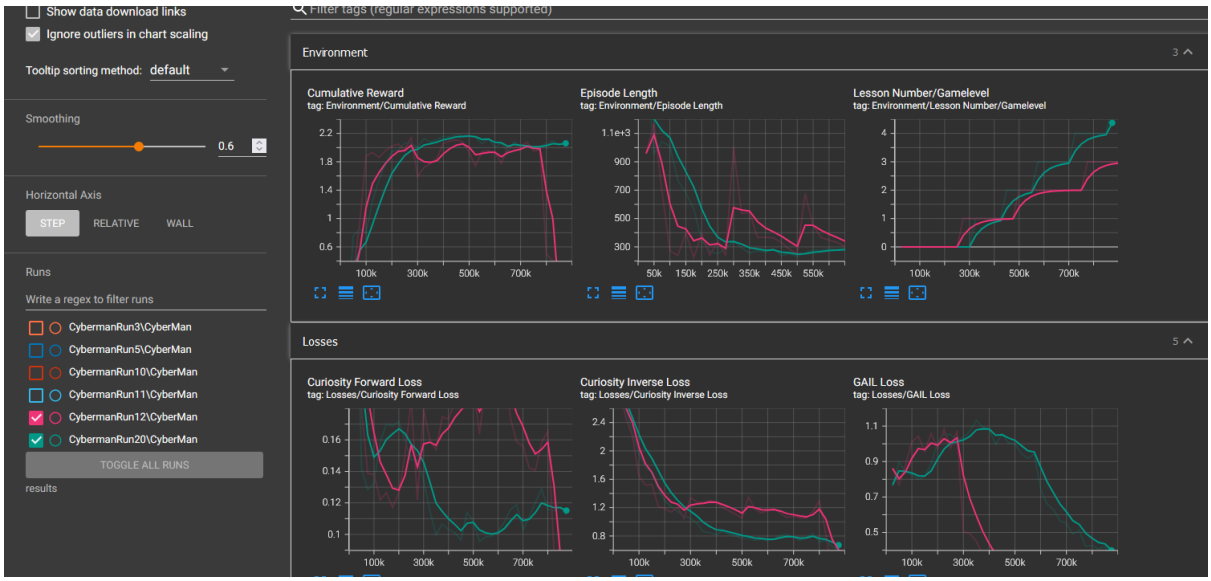


Figure 1 Fast Early Training

However the Dalek is picking up pace, and at Level 9 will probably be too much of a challenge, as the Agent will remain within the central arena, and the Dalek will always see, and catch the agent at this level. However Transition to level 9 at around 1.7 Million Steps. The Agent seems to be coping reasonably well, as Random Positions still generally to the South, so still just enough time, typically, 70% to avoid encountering the Dalek. However there is a very poor perception of the Dalek.

Figure 2 Level 9, First drop in performance Challenge

Fails to get Beyond Level 9.

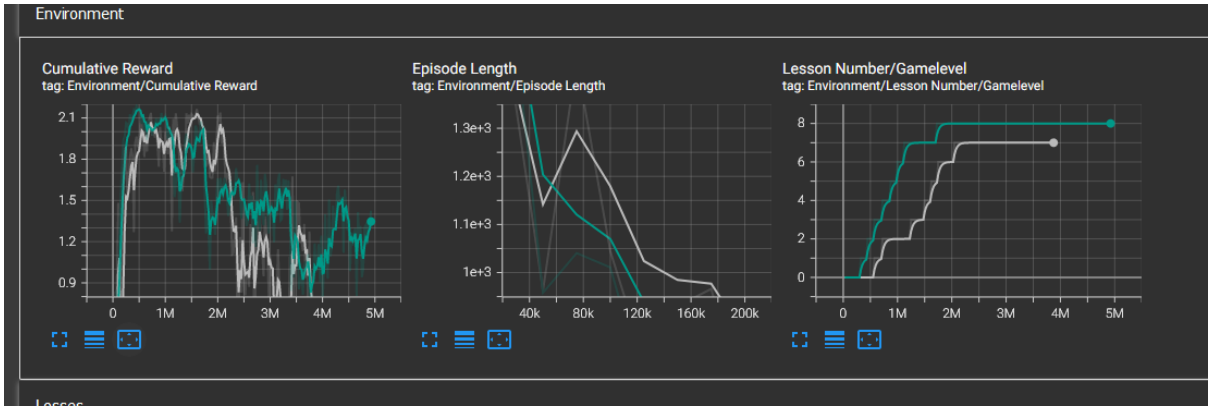
Run 21

The Dalek Perception Problem

Even though Agent sees the Dalek, it is often at distance, obscured by closer objects of Interest. So the perception of the Dalek is rather weak. Which suggest need an Explicit Distance to Dalek Observation, to help appreciate danger.

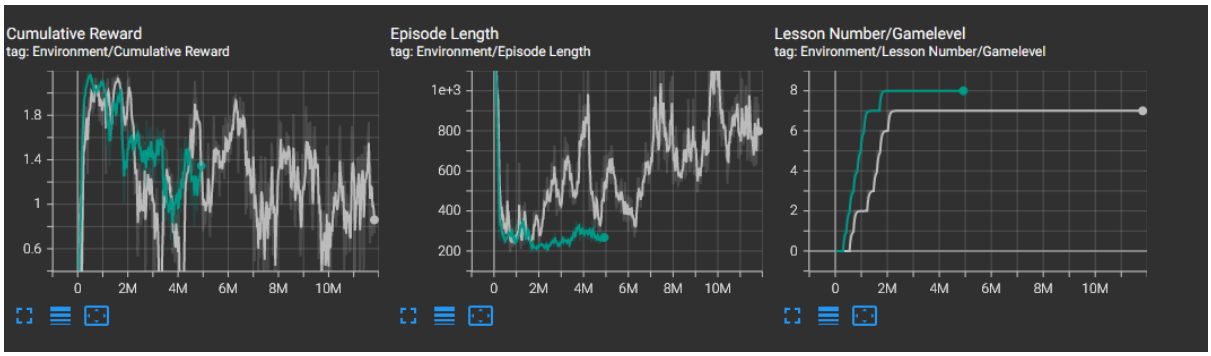
- Perhaps conditional Observation. If cannot see Dalek Distance == 5.0f, otherwise IF can see – then return the Actual Normalised Dalek Distance.
- Also Noted that was Missing the Bool Indicator, Being Seen, as a valuable Threat Observation
- Clear down the Game scores, upon any change of Levels
- Increase in Dalek speed for Levels 1..4, but Decrease in Dalek Speed for level 7..10.

The Agent now seems slower to learn, perhaps a more complex Observations. But it looks as though it is now struggling on Level 8.



Where Green is previous Run 20, Grey is Run 21.

Need to see if eventually gets past Level 8, suspect from past experience runs it will not ☹



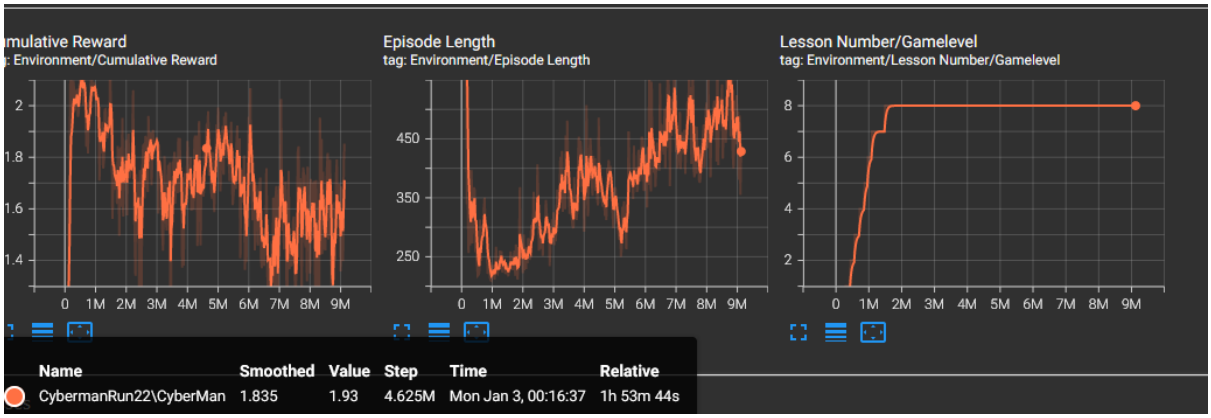
Seems to be dumping close, but not precise onto Green platform. It knows it has to pace at green, but drops too early/ far.

Run 22

Next Suggestions –

- Avoid SW Wall 3 being a Trap, shorten to allow Gap at Bottom. – AND **Rebake all Environments**
- Make Goal a bit easier again, at 1.25 x Threshold
- Reduce Curiosity Strength down to 0.02

Gets to level 9, in 2 Million Steps

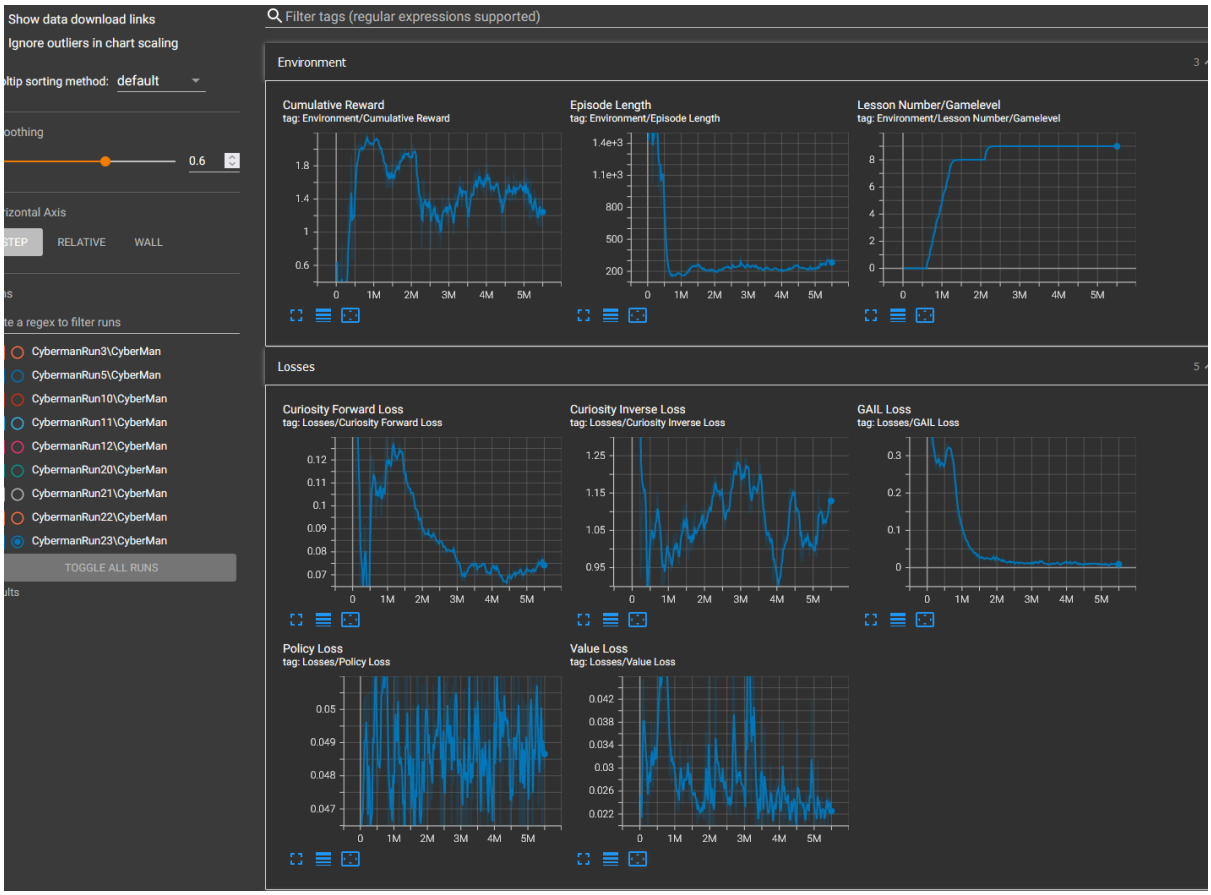


Still very little concept of the Dalek as a bad thing. Just continues its task, with little perception or regard to the Dalek !

Run 23

- Use Stacked Frames x 4 – so as to hopefully get a concept of the Dalek Motion coming Towards as a Threat
- Add an explicit Observation – Distance to Dalek (In addition to Relative x ,y)
- Set MaxSteps to 20 Million

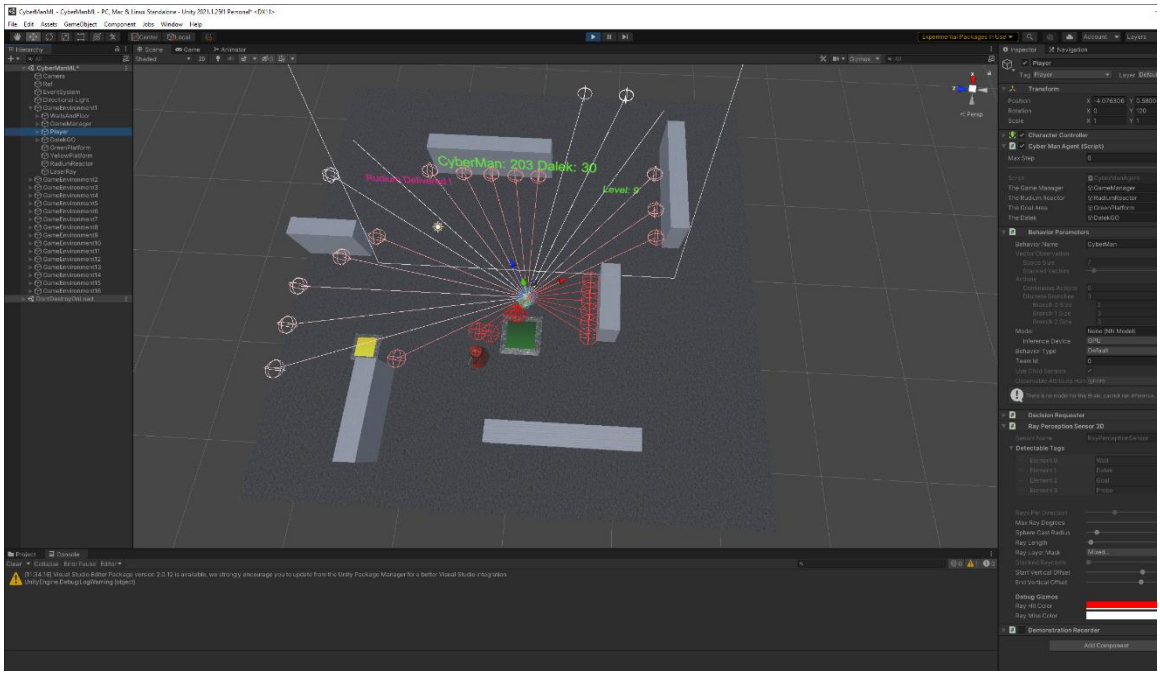
Now Gets to level 10 - Then pretty much stuck – Note that a Dalek speed at 0.5f always has caused the Inability to compete (previously was level 5 originally, then 7, then 9)



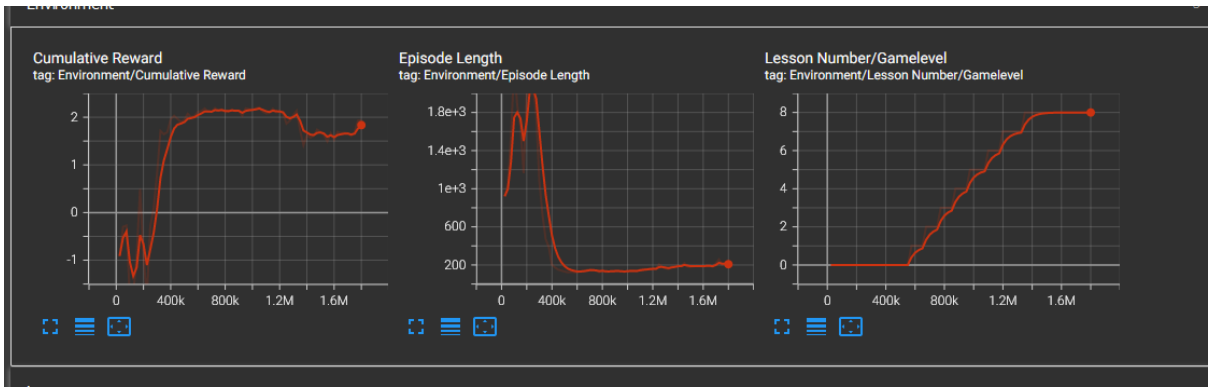
Looking like finding easy to climb up levels 1 to 8 within a Million and a Half steps

Run 24:

- Steeper Negative Step Reward Gradient, whilst being Seen by Dalek.
- -0.2f Reward for any Place down. (Increase successful Place down to +2.2f) To Penalise excessive place downs
- Change Dalek See Distance to ~1.75 x Current Laser Range Distance
- Change Dalek Speed profile to get to the troublesome 0.5f speed level (0.5f Still at level 10), but increase Dalek speeds at lower Levels. So level 9 is now more of a challenge with a Dalek speed of 0.4f
- Change Number of Episodes at Level Threshold Reward to 750, to slow more consistent Level Growth.
- Increase Fan of Ray Perception, to 140, with 140 rays, so a Ray every 10 degrees. So the Agent has more perception to the sides, and a little behind.



Run 24 Quickly gets to the (now more Challenging, 0.4f speed) Level 9 within 1.5 Million Steps.



But Now gets stuck at level 9

Suggestions

Watching Sebastians Shooter Video (Unity ML-Agents 1.0 - A.I. Shooting Game (FULL WALKTHROUGH) - YouTube)

- Suggests that the Time_Horizon Needs to be Long enough to capture Trajectories of relevance (Reward Trajectories) So maybe our Time_Horizon needs to be longer ? However a more authorative correct interpretation is
  "Horizon, in simple words, is how far in time steps a present reward can be associated with a past action (or a past action can be accounted for in the present reward)"
- Sebastian shooter Agent has two sets of Observational Rays.
  a) Precise Front Facing, with around 20 Rays in 80 degrees front (No Spheres, just Lots of Rays)
  b) Lower Precision All around, Around 14 All way Round.

Danielis Golubovskis MSc

Note Found a reference to an MSc implementation of Hide and Seek, using ML Agents. Inspired by the familiar OpenAI ~~videos. Danielis Golubovskis: MrpHDanny (Daniel G) (github.com) with the specific Repository containing his Word dissertation : MrpHDanny/Undergraduate-Dissertation: Replacing traditional AI in video games using Machine Learning Agents (github.com)~~

His proposed Yaml configuration:

- Batch Size 128
- ~~Buffer size 2048~~
- ~~Hidden Units 128, But with num_Layers =3.~~
- ~~Beta 5.0e-2~~
- ~~Normalize = true~~
- ~~Time_horizon 64~~

Note Danilleis did not use either GAIL or Curiosity Networks

~~**Run 25:**~~

~~**Make a significant changes to the YAML configuration**~~

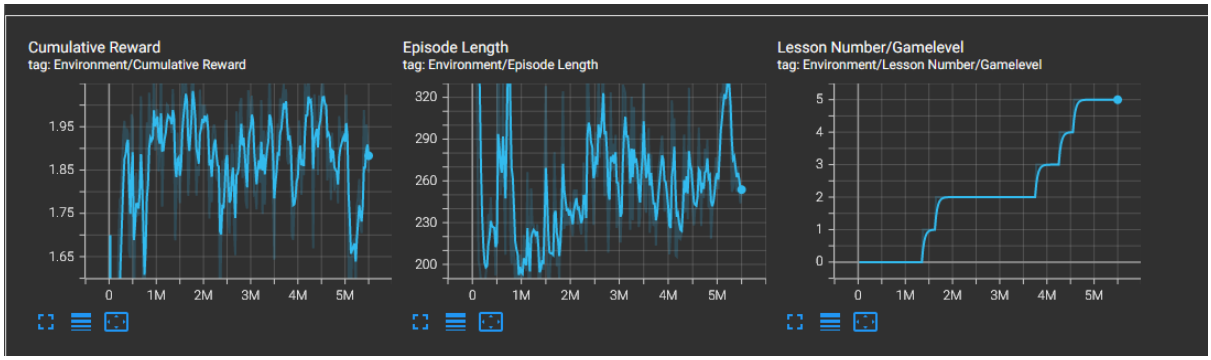
- ~~Reduce Batch size down to 128 (from 256)~~
- ~~Reduce Buffer size down to 2048 (from 5120)~~
- ~~Increase beta to 0.025~~
- ~~Normalise Set to **true** (from false)~~
- ~~Change Network to hidden_units 128, but num_layers = 3~~
- ~~Reduce Time_horizon down to 64 (from 256)~~
- ~~**Remove GAIL AND Curiosity networks !**~~

~~Expect Now to take a Long Time to learn Initial steps Level 1 Objectives without GAIL~~

May need to recover some of these back ! – lets see how this radical change impacts behaviour **(See Later)**

No changes to sensors (as Danielles appears to have experimented with Raycasts + Vectors)

Yes Taking a rather longer to Train and more compute. With deeper network. 5 Million Steps to get to level 6. Gets to Level 9 at 9 Million steps.



Again struggling at Level 9 (now with Dalek speed 0.4f) -Despite the proportion of evasive encounters only being around 20% (around 50% are straightforward, south of Arena scenarios) . Will need a lot more Training to meet more encounters. **May need to Train Towards 40 Million !**

Notice some slight avoidance of Dalek Tactics some of the time, certainly on Goal Side, but not so much on Yellow Pick Up approach.

~~**Run 26:**~~

- Instead of random vector LERP on Dalek initial positions, Just Random Dalek in NE or NW points, to either side of Walls (c.f. the V2, V3), and Initialise the first Waypoint, on Episode Begin/ Reset to being centre of map (0,0) so the Dalek always approaches the centre of the Map. So as to get more frequent exposures to the Dalek.
- ~~Reducing Decision Requester Period to 4~~
- ~~Increase the learning Rate to 0.0003 (from 0.0002)~~
- ~~Decreased beta back to 0.02~~
- Max Steps increased to 40 Million.

Stuck at level 1 !!

~~**Run 27:**~~

- ~~Reduce the learning Rate down to 0.00015~~
- ~~Set beta back to 0.025~~

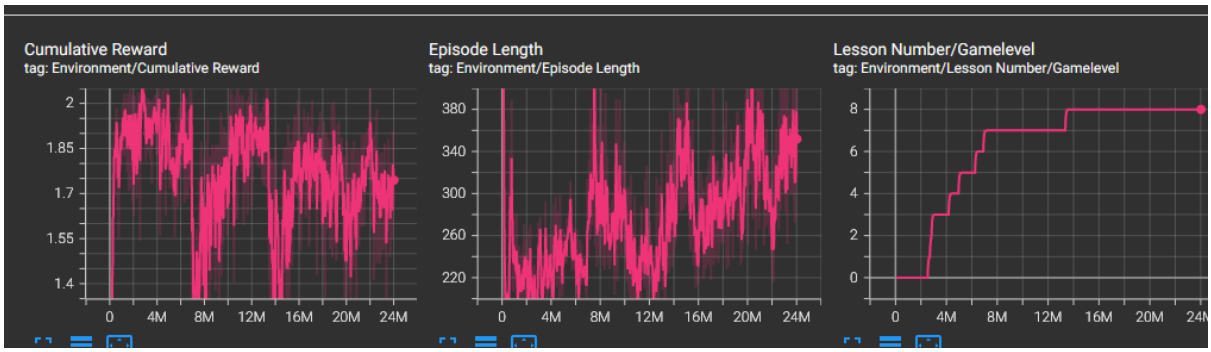
~~**Still not getting past Level 1**~~

Run 28

- Revert Decision Interval Back to 5
- Rays per Direction back to 14 (Every 10 degrees)
- Increase Time_Horizon to 128
- Change Buffer Size to 2560

- Reduce Max_Steps to 30 Million (as profile Learning rate on that)

Now gets to Level 9 Again, rather slow to get to level 9. but still not properly evading the Dalek, and still stuck at level 9 at 24 Million.



Some evasion of the Dalek, but does typically get caught

Run 29

- Maybe additional Forward Rays sensors, (Elevated) To see Dalek – Above Probe and Goal Colliders. So Dalek does not get Obscured by those two objects, when going towards.
- Need to reduce the height of the Goal and Probe colliders.
- Suspect the Dalek Laser is also being Obscured, so the Dalek is not a threat if either Goal or Probe is in same line of sight! : Need to Raise the Dalek Head sight above Colliders.
- Return Learning rate to 0.0002
- Increase beta to 0.04

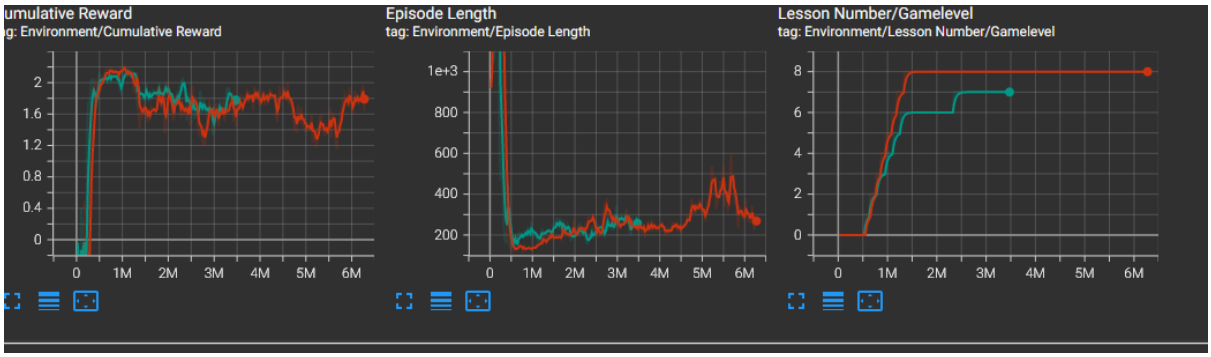
Fails to get out of Level 1 !

Run 30

Revert to Run 24 Hyper parameter configurations – as was the fastest learner (Simpler 2 layer network, normalise false, reinstate GAIL etc)

- GAIL and re recorded demo put back in ?
- Albeit Time Horizon left at just 128, not 256

At least now gets to level 8 again in 3 Million Steps (red Run 24 stuck at level 9, Green Run 30)



Still significant proportion of Probe and Goal spawns, to the south making it relative easy wrt to Dalek on Level 8.

Gets to level 9, in 7 Million Steps and Level 10 at 13 Million. Does learn to avoid, and to some extent evade the Dalek.

It Now gets to 11 at 20 Million Steps!

So with now random spawn positions, and the agent has a much bigger challenge to discover find the Probe and the Goals around the arena. No spatial memory, and has not learn't to really explore effectively. Just rather Random and Avoidance.

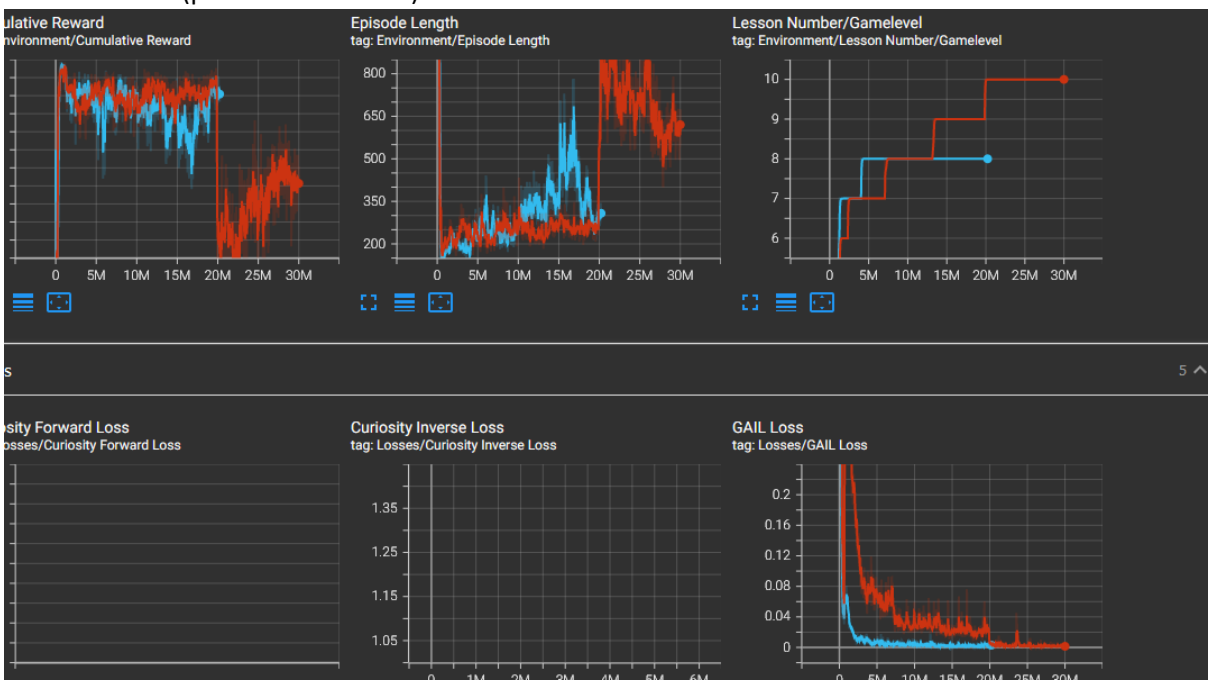
Average Reward Profile has dropped off, so Suspect will not be able to get out of Level 11, within 30 Million Episodes

Run 31

- Clear Narrative at Player step == 200
- Fairer to give the Agent longer Ray length on the General Search. (from 30 to 40.0f) So easier to see Goal and Probe at Distance. Slightly wider fan out to 140 degrees for Dalek rays, and 120 for general Probe/ Goal
- Run for 40 Million Steps
- New Demo Recoring31 at ~~Game level 8~~ (as resolution of ray perceptions have changed)
- This Level8 recording looks as though the initial Training was very excessive exploration, and little no direct routes. Unlikely to get out of Level 1
- Record at Game Level 6, with very short routes.

Slower struggle to get beyond Level 9, as still at level 9 at 20 Million. Although performing and avoiding behaviour reasonably well.

Run 31 in Blue (previous 30 in Red)



Run 32

- Reduce learning rate to 0.00015. (as it is now a longer run)
- Increase Buffer size to 5120 (Alignment with Run 23/24)
- Add Curiosity Back in (Alignment with Runs 23/24)
- Increase Time_horizon 256 (Alignment with Runs 23/24)

Noting we now have curriculum lessons step up at 750 not 500 as previous in R23/24.

Climbs up to Level 10 in within 8 Million Steps.

But Cannot seem to get to Level 11. Certainly not in 20 Million Steps.

Run 33

- Double learning Rate 0.0003 (from 0.00015)
- Increase Time Horizon further to 512 (from 256)
- Increase Buffer_Size further to 10240 (from 5120)
- Reduce beta from 0.005 to 0.004

Again gets up to level 8 pretty quickly, in about 1.2 Million Training episodes.

But having more problems at getting up from level 8. But similar to run 30, which took to 7 Million before getting to level 9 (and then on towards level 11 at 20 Million)

Struggling at level 8, so abort at 8 Million

~~**Run 34 Suggestions**~~

- ~~Reduce Batch size to 128 (As just a few discrete Actions)~~
- ~~Increase Buffer Size again to 20480 (*MrMonday advice)*~~
- ~~Reduce Learning rate back to 0.0002~~
- ~~Increase main Gamma to 0.9975 (*MrMonday advice)*~~
- ~~Change beta Back to 0.005~~
- ~~Time_horizon increase even further to 1024~~

~~***Longer processing stalls, as now doing a lot more processing (I guess time Horizon), so will take longer clock time***~~

Only got to Level 2, so revert back

~~**Run 35**~~

~~Revert back some settings~~

- ~~Return Batch Size to 256~~
- ~~Gamma to 0.995~~
- ~~Reduced time_horizon back down to 256. (as per runs 23/24 and 30)~~

~~*So Gamma and Buffer sizes still higher than many previous runs (from MrMonday)*~~

Stuck at Level 2 again !

Run 40

Revert most settings towards runs 30/31

- Batch Size 256
- Buffer size 5120 (not 2560)
- Gamma 0.99
- Time_Horizon 128
- Learning Rate 0.0002
- Beta 0.005
- Max Steps 30 Million (**not 20 Million**)
- ***But with Curiosity still left in and***

Rapid rise to level 8, then remains stuck at level 8 at 5.5 Million

Run 50

Major overhaul in line with Dragon Environment

- ⇨ Reduce Ray Casts.
  - General Goal/ Probe sensor Reduce to 8 Rays across 90 degrees
  - Dalek Sensor Reduce to 12 Rays across 120 degrees
- ⇨ Flatten the Action Space to size 6: 6: NOOP, Fwd, Left, Right, Pickup, Place
- ⇨ Add Booleans as direct Observations
- ⇨ Add Dalek Relative Position – regardless of can see
- ⇨ Configuration changes
  - Sig Increase the Batch size 1024
  - Increase Buffer_size 10240
  - Beta 0.01
  - Learning Schedule **constant**
  - Reduce Time_horizon down to 64
  - Curiosity strength reduce to 0.01
  - Reduce GAIL strength to 0.075
- ⇨ Re Record Demo for GAIL

So differenced to dungeon are we still have curiosity, our network is still 512, learning rate is lower at 0.0002, and max steps is 30 Million, not 20

Stuck at level 2 ! - Not getting anywhere !

Run 51

- ⇨ Remove curiosity
- ⇨ Reduce buffer size back down to 5120
- ⇨ Batch Size 512
- ⇨ Increase Learning rate to 0.0003
- ⇨ Reduce beta to 0.005
- ⇨ Increase GAIL strength back to 0.1

So more similar to dungeon, half of its Buffer and Batch size, but same Time_horizon, learning are and No curiosity.

Seems to be struggling at Level 3, even at 3 Million Steps

Run 52

Possible go back toward s Near to Danieles MSc [See Run 25] (apart from the Normalise, and deeper Network)

- ⇨ Reduce Batch size to 256, and Buffer down to 2048, (c.f. Run 30)
- ⇨ Increase Beta to 0.01
- ⇨ Increase Time_Horizon back to 128 (c.f. run 30)
- ⇨ Learning rate slight reduce to 0.00025
- ⇨ ***Remove GAIL see if it learns any basic levels***

Likely to struggle al lot to climb first levels so need to give it some 4 million steps

Gets Stuck on an Action sequence with No Movement, Action. Does not seem to shift away.

Run 53

- ⇨ Re Profile ALL Rewards to accumulate to -1.0 to +1.0f (use **SetReward** for Mission Complete and Deaths)
- ⇨ So Change Curriculum Step Up Reward Threshold of 0.85f
- ⇨ Reduce size of Main Policy Network (hidden nodes) from 512 down to 256.
- ⇨ Put GAIL back in and Re capture a demo at very simple Level 3.

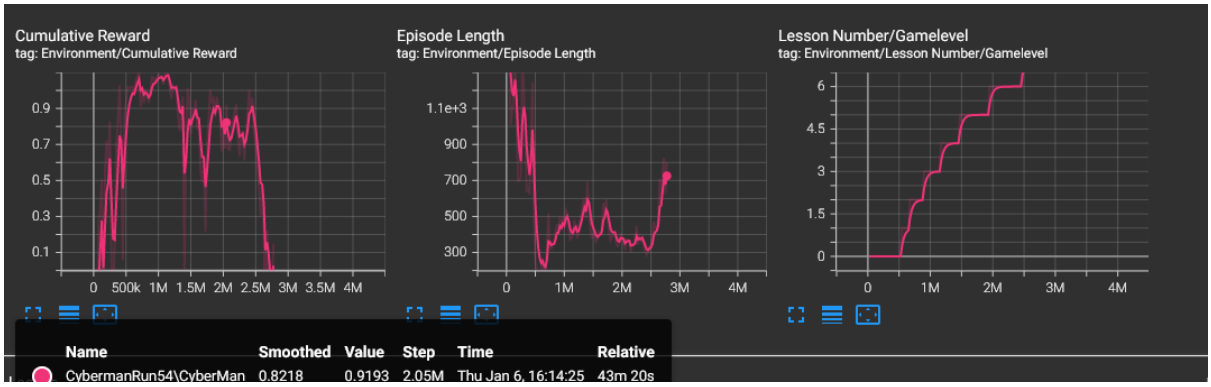
Gets to level 5 in 1.5 Million Steps
Level 5, the Dalek now starts on its attacks, so suspect will remain on this level for some time for it to properly assess

Could not get beyond Level 5 – The Dalek is more aggressive – However an Error in the YAML, had a threshold left at 2.0, so could not promote to level 6.

Run 54:

- ⇨ Correct Reward Threshold error
- ⇨ Reduce Leaning rate to 0.0002

This run Gets to Level 8 in 2.5 Million Training Steps



Stuck at level 8

Run 55:

Running out of Episodes, needs to be better policy than being killed off by Dalek. At least perhaps learn to avoid Dalek, and move away at the very least.

- ⇨ So Set Run out of Episodes => -0.75f, and those walk penalties max at -0.75f/max steps
- ⇨ Reduce Beta to 0.005

Level 8 seems to be a big step, and struggle (previously it was when Dalek speed =0.5f)

This is the current Dalek Speed Profile:

```
if (GameLevel < 3) TheNavMeshAgent.speed = 0.0f;
if (GameLevel == 3) TheNavMeshAgent.speed = 0.05f;
if (GameLevel == 4) TheNavMeshAgent.speed = 0.1f;
if (GameLevel == 5) TheNavMeshAgent.speed = 0.15f;
if (GameLevel == 6) TheNavMeshAgent.speed = 0.2f;
if (GameLevel == 7) TheNavMeshAgent.speed = 0.25f;
if (GameLevel == 8) TheNavMeshAgent.speed = 0.3f;
if (GameLevel == 9) TheNavMeshAgent.speed = 0.4f;
if (GameLevel >= 10) TheNavMeshAgent.speed = 0.5f;    // 0.5f remai

LaserDistance = 7.5f;
if (GameLevel < 5)
{
    LaserDistance = 7.5f;
    MaxSeeDistance = 12.5f;
}
if ((GameLevel >= 5) && (GameLevel <= 10))
{
    LaserDistance = 10.0f;
    MaxSeeDistance = 17.5f;
```

Run 56:

The Step Speed from 0.25f => 0.3f seems to be causing a problem. Change Dalek speed profile*:*

Level	*Prev*	*New*
<3	0.0	0.0
3	0.05	0.1
4	0.1	0.2
5	0.15	0.25
6	0.2	0.275
7	0.25	0.3
8	0.3	0.35
9	0.4	0.4

= 10	0.5	0.45
>10	0.5	0.5

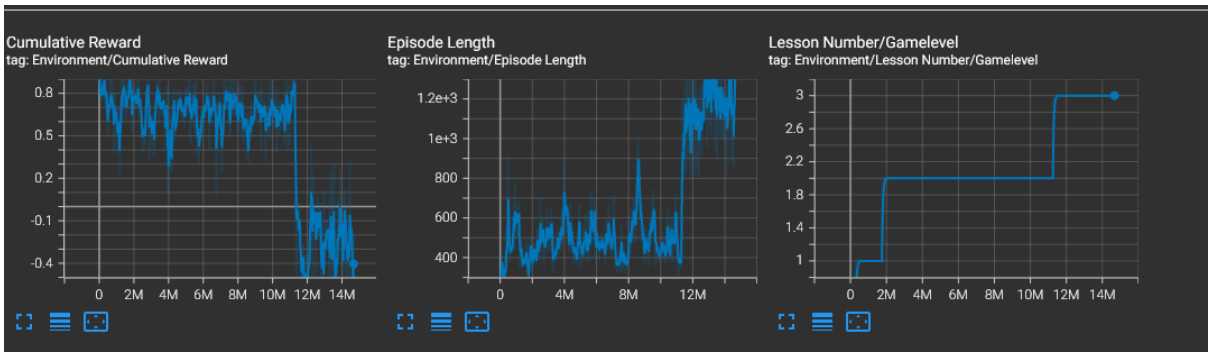
This may make Level 7 the new challenge point

Got up to Level 10 at 28 Million Training Steps
Reasonable but not great behaviours. Needs further Training perhaps

Run 57
Can we continue, using the Brain 56 ? using `--initialize-from=<run-identifier>`

- ⇨ Change the YAML, to Start at Level 8
- ⇨ Start Training Run 57 with `--initialize-from=` CybermanRun56
- ⇨ Set Max Training Steps to 50 Million

Restarted Training at Level 8, quickly gets to level 9, then back t level 10 at 2 Million, and then onto Level 11 at around 11.5 Million steps. Level 11 major drop in average rewards, maybe difficult to through Level 11, now with its Random Spawn Points.



Further Suggestions

Run 58:

Continue Brain at Level 9:

Change Reward Profile, to ensure that walk penalty is properly recognised:

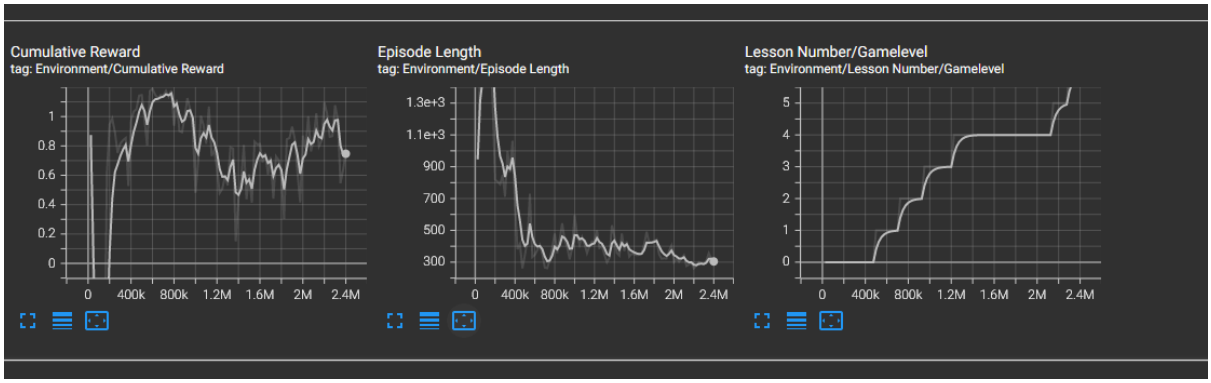
- ⇨ ***ADD 1.25f*** *for successful Place down, so that do recognise penalty for excessive movements, and appreciate rewards for Pickup and Place down.* ***And Change first Pickup +0.2f***
- ⇨ *So Pickup and place down will cancel each other out (+0.2f and -0.2f for a successful single pickup and place down Trajectory.)*
- ⇨ ***Change movement penalty*** *– General: -0.3f/Maxsteps, Can See:-0.5f/Maxsteps, -0.15f/Maxsteps if have the Probe*
- ⇨ *Set the Change Level Reward Threshold to 0.825f*
- ⇨ *Reduce Dalek Speed at level 11, to compensate for more diverse Spawn Locations*
- ⇨ *Reprofile Levels 12,13,14 Dalek Speeds*

Run 58 Gets to level 11, but still just Struggles with the random placements. Little to no Improvements after another 20 Million Steps

Backup and Re think !

Run 60

- Create more sensor coverage behind Agent (General Probes, Goal) Add more Rays/ Rear Range to the Ray Sensors. Set to 160 degrees, 16 rays coverage for both sensors
- Reintroduce Relative Probe/ Goal positions Observations Relative x,y, regardless.
- Add Delta Dalek Positional Change Delta X, DeltaX since last Collection– in lieu of its relative Velocity.
- Remove Stacked Frames
- Step Penalty: Penalise being Idle (0.5f), more than Walking (0.25f) to hopefully encourage more movement, and hence exploration
- Add NavMesh Obstacle to the Probe Object, so as Obstruct the Dalek (Carve =True)
- Increase Hidden Nodes back to 512. (But it does look as though a 256 node does have the capacity for learning the basic pickup and place down job, and for some Dalek avoidance)
- Reduce the MaxEpisode Steps within Agent to 10 10,000
- ~~Add a very small amount of curiosity back in~~
- ***Re Demo and Start Training from level 1 Again***



Seems to be learning the Basics, getting to level 5 at 1.5 Million, and then onto level 7 within 2.5 Million, and then Level 9 at 5 Million.

Does not seem to be avoiding the Dalek Yet.

Noted that the Agent Rays are not Detecting the Dalek. Because the Collider was removed from the Dalek. Ray cast sensors cannot see The Nav Mesh Agent without The Collider.

Run 61:

- Add The Collider Back in and attempt to Resume Training from level 8.

Reaches Level 11, 2 Million after the restart.

Although The Agent seems more mobile, and exploratory moves than before, the Dalek seems to be having an easier time in getting to the Agent.

Great ! Get to Level 11, pretty quickly in around 5 Million Steps overall. *But in level 11, its is again rather Stuck and has little or no concept on how to explore/ find the Probe/ Goal when it cannot see them.*

Run 62:

- Use Bearing (wrt to Player Forward) and Range Observations towards Goal, Probe, rather than relative x,y ? . These Bearing, Range observations should be easier to map onto Rotate left, Right Actions.
- Use **Vector3.SignedAngle** From(Unity - Scripting API: Vector3.SignedAngle (unity3d.com) For example:

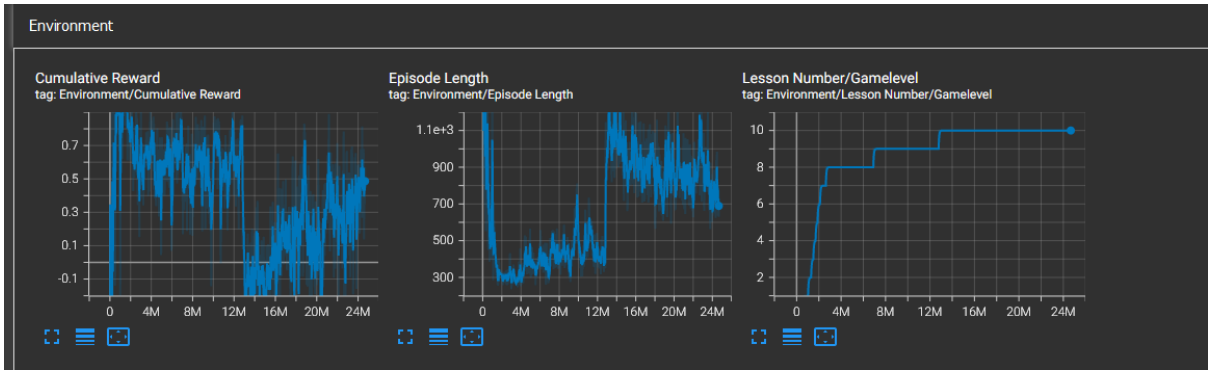
```
Vector3 targetDir = target.position - transform.position;
        Vector3 forward = transform.forward;
        float angle = Vector3.SignedAngle(targetDir, forward, Vector3.up);
        if (angle < -5.0F)
            print("turn left");
        else if (angle > 5.0F)
            print("turn right");
        else
            print("forward");
```

- Then use the **RelDirectionTo (Probe) and (Goal) and Dalek** and **Vector3.distance/50.0f** to those three, instead of Relative x, z positions. So these will act similar to, but as explicit (b, r) observations.

- And just use ***closing Dalek Speed Delta Vector3.Distance differences*** rather than Delta X, Delta Z observations, by storing ***PrevDalekDistance.***

  ClosingSpeed = (CurrentDistance – PrevDalekDistance)/~~Time.deltaTime~~
- *Checked the typical magnitude of this value, and scale up towards 1.0 range by x 25.0f.*
- Re Demoed the new Sensors into DemoRun62.demo

Note the closing Speed may be too aggressive, since the player motion towards to Dalek will be very Fast signal to that vector



Climbs up to Level 9 in first 3 Million Steps, and gets to level 11 at 13 Million steps.

Run 63

- Continue to Run `--initialize-from=Run62.`
- For Another 10 Million Steps, but with a Linear Decay

Used This Run as source of You Tube and GitHub presentations

Further Suggestions

Rethink the Environment with ramps and steps.

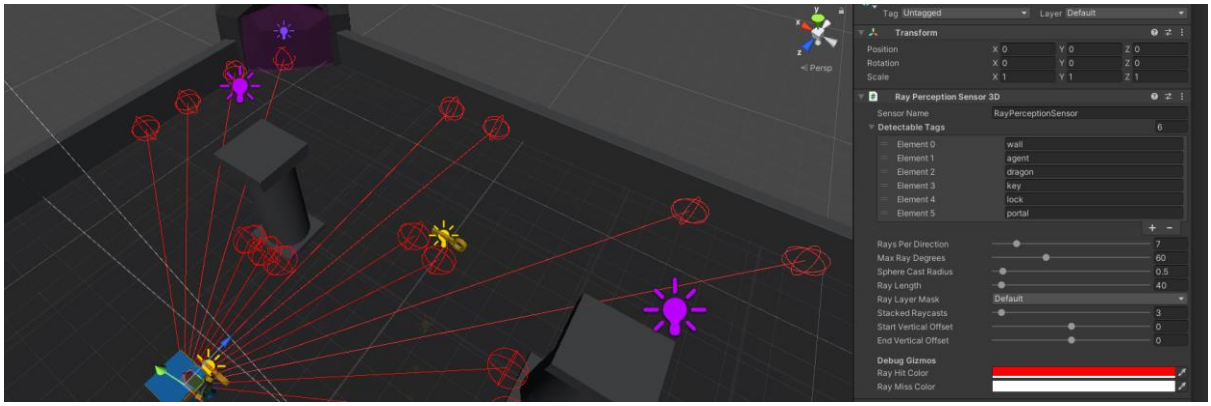
<u>Dragon Dungeon Review</u>:

- Batch_size 1024 - so really large
- Buffer size 10240 - so really large
- Learning rate 0.0003
- Beta 0.01
- **Learning Rate constant**
- Hidden units: 256 and num layers 2
- Gamma 0.99
- Normalise false
- **Time_horizon 64 - so pretty small**
- Max Steps set to 20 Million

No curiosity or GAIL. ***Dragon Dungeon suggested to take only around 2.4 Million Training steps Train without any GAIL***

Dungeon Agent uses mainly Ray cast, with Spheres (0.5radius), but only 7 Rays, across 60 degrees forward. All items Wall, Dragon, Key,portal etc tagged to detect)
+ 1 explicit Observation [Does have Key) - This is added simply as **a direct Boolean**
Sensor.AddObservation(IHaveKey) // where IHaveKey is a boolean

⇨ ***Agent Can use fewer Rays, with a narrow field of view***



Uses a **Flat Action Branch size 7**: Just simple Move agent (Fwd,Back,Rotate Left, Rotate Right, Translate Left, Translate Right) , so Basically NOOP + 6 basic Move Actions

```
11 references
public override void OnEpisodeBegin()
{
    MyKey.SetActive(false);
    IHaveAKey = false;
}

8 references
public override void CollectObservations(VectorSensor sensor)
{
    sensor.AddObservation(IHaveAKey);
}

/// <summary>
/// Moves the agent according to the selected action.
/// </summary>
1 reference
public void MoveAgent(ActionSegment<int> act)
{
    var dirToGo = Vector3.zero;
    var rotateDir = Vector3.zero;

    var action = act[0];

    switch (action)
    {
        case 1:
            dirToGo = transform.forward * 1f;
            break;
        case 2:
            dirToGo = transform.forward * -1f;
            break;
        case 3:
            rotateDir = transform.up * 1f;
            break;
        case 4:
            rotateDir = transform.up * -1f;
            break;
        case 5:
            dirToGo = transform.right * -0.75f;
            break;
        case 6:
```

Note a Linear decay is recommended for PPO. This will erode Learning rate towards a Zero Learning Rate at max Training Steps. Note this Linear decay is also applied to Beta and Epsilon.

- ~~Perform GAIL Recordings at level 9~~
- Maybe attempt to Flatten the Actions to a single Branch Action size 6: NOOP, Fwd, Left, Right, Pickup, Place.
- More Bias, influence of Probe and Goal to be spawned in the North after Level 5. Perhaps use new V1..V4 vectors. So as to encourage more threat and evasion
- Maybe increase Dalek Laser range at level 6 and 7

~~Maybe try reducing Decision Requester Period to say 2 perhaps ? (Note this will slow down the Training computation considerably)~~

Nav Mesh Interactions with other Objects (E.g. Dalek nav mesh Agent.)

From Unity - Manual: Using NavMesh Agent with Other Components (unity3d.com)

- Need to Add a Collider component
- Add a Rigid Body Component. (But Turn ON Kinematics, if. Physics off)

NavMesh (Static) is baked set of surfaces and objects to define walkable routes. Based upon all Game Objects marked as Navigation Static

Use **NavMeshObstacle** component for Moveable Objects to that the Agent will Avoid (None walkable) e.g. a Barrel or Crate. E.g. for the Rudium Probe