# Training management

## Runs:

BaseEnvironment-1.2.11-0: Speed: 2, DI: 5, TS: 60 Training time: 1650 seconds

BaseEnvironment-1.2.12-0: Speed: 1, DI: 5, TS: 60 Training time: 1500 seconds

BaseEnvironment-1.2.13-0: Speed: 1, DI: 3, TS: 60 Training time: 1350 seconds

BaseEnvironment-1.2.14-0: Speed: 1, DI: 7, TS: 60 Training time: 1750 seconds

BaseEnvironment-1.2.15-0: Speed: 1, DI: 10, TS: 60 Training time: 1800 seconds

BaseEnvironment-1.2.16-0: Speed: 1, DI: 5, TS: 100 Training time: 1350 seconds

Mass: 25, drag: 2

BaseEnvironment-1.2.17-0: Speed: 1, DI: 5, TS: 20 Training time: 2020 seconds

BaseEnvironment-1.2.16-1: Speed: 1, DI: 5, TS: 100 Training time: 1563 seconds

Mass: 1, drag: 0

BaseEnvironment-1.2.16-2: Speed: 1, DI: 5, TS: 100 Training time: 1360 seconds

Mass: 50, drag: 4

BaseEnvironment-1.2.16-3: Speed: 1, DI: 5, TS: 100 Training time: 1390 seconds

Mass: 1, drag: 1

CLStaticDynamicObstacles-1.0: Preferred set-up, no bound of max. Steps.

CLStaticDynamicObstacles-1.1: Preferred set-up, bound of max. Steps: 1000.

CLStaticDynamicObstacles.1.0.1: Preferred set-up, no bound – just another run of 1.0, to get diversified look at the behaviour.

CLStaticDynamicObstacles-1.1: Preferred set-up, no bound – just another run of 1.0, to get diversified look at the behaviour.

CLStaticDynamicObstacles.1.5: Preferred set-up, bound of max. Steps: 4000. Alternative counts from CollectOobservations().

CLStaticDynamicObstacles-1.6: Preferred set-up, bound of max. Steps: 4000. Alternative counts from AgentAction().

CLStaticDynamicObstacles-1.7: Preferred set-up, bound of max. Steps: 4000. Alternative counts from: AgentAction(), Update() and FixedUpdate().

**1.7: shows that the customer counter should be implemented in FixedUpdate() in order to reach the max step count.**

CLStaticDynamicObstacles-1.8: Preferred set-up, bound of max. Steps: 4000. Alternative counts from: AgentAction() and FixedUpdate().

**1.8: Was a test to see if and how much over the max step the counter went, now that 1.7 showed that we reached the max step. The counter went precisely 5 steps over, which is equivalent to the last action is being taken as the 4000-step-threshold is reached.**

**Bonus insight: The max-step could also have been reached from the agent action alone.**

CLStaticDynamicObstacles.1.0.3: Default hyperparameters, no bound – TS: 50, Training time: 2450

CLStaticDynamicObstacles.1.0.4: Default hyperparameters, no bound – TS: 100, Training time: 2550

CLStaticDynamicObstacles.1.0.5: Default hyperparameters, no bound – TS: 50, DistToTarget included in VectorObs; Training time: 3450

**Concurrent training:**

**CcCLStaticDynamic-1.0: All four environments in action from start to end. Only aggregated stats.**

**CcCLStaticDynamic-1.1/2/3/4: Increasing number of environments in action (2,2,2,3,4,4). (2) does not have separate stats. For each of the environment and so only aggregated stats.**

**CcCLStaticDynamic-1.5: All four environments in action from start to end. Separate stats.**

**CcCLStaticDynamic-1.6: like 1.5 but with four times higher buffer.**

**CcCLStaticDynamic-1.7: All four environments in action from start to end. Separate stats.**

**Vector stacking:**

Building off CLStaticDynamicObstacles-1.7/8.

CLStaticDynamicVS3: Shows good results

CLStaticDynamicVS6: Catastrophic results.

~~CLStaticDynamicVS9~~

Reward Shaping:

RewardShaping-1.0: Unbounded, 0.0005 reward to getting closer to the target.

RewardShaping-1.1: Unbounded, 0.0005 reward to getting closer to the target.

RewardShaping-1.2: MS: 4000, 0.0005 reward to getting closer to the target.

**Visual agent:**

VisualAgent-1.5: Default hyperparameters, unbounded. Grey scale. OS: 1. TS: 50, Training time: 11000

~~VisualAgent-1.0: Default hyperparameters, unbounded. Colour. OS: 1. TS: 50, Training time:~~**~~Interrupted because of too long dropouts.~~**

~~VisualAgent-1.1: Default hyperparameters, unbounded. Colour. OS: 1. TS: 50, Training time:   
Lowered buffer to 5120 from 10240, Batch unchanged.   
Still dropout but decreased.~~

VisualAgent-1.2: Default hyperparameters, unbounded. Grey. OS: 1. TS: 50, Training time: 7265  
Buffer: 2560, Batch 512.  
Still dropout but decreased to 40 seconds.

VisualAgent-1.3: Default hyperparameters, unbounded. Colour. OS: 1. TS: 50, Training time: 8870  
Buffer: 2560, Batch 512.  
Still dropout of around 40 seconds

VisualAgent-1.4: Default hyperparameters, unbounded. Colour. OS: 1. TS: 50, Training time: 8600   
Buffer: 2560, Batch: 512. NumLayers decreased to 1 from 2 and NumEpoch decreased to 2 from 3.

**Full Set-up:**

*Curriculum 1.0: Static -> Dynamic -> Semi-dynamic*

*Curriculum 2.0: Static -> Semi-dynamic -> Dynamic*

FullSetUpCertain-1.0: Default hyperparameters, bounded; MS: 4000. VectorObs Only. OS: 3. TS: 50. Curriculum: 1.0, Steps: 1.000.000, Training time: 8000

FullSetUpCertain-1.1: Default hyperparameters, bounded; MS: 4000. VectorObs Only. OS: 3. TS: 50. Curriculum: 2.0, Steps: 1.000.000, Training time: 8500

**Detour:**

***FullSetUpCertain-1.1.1:***

Default hyperparameters, bounded; MS: 4000.   
VectorObs Only – four additional observations added (Dens, dist) on the CrowdedAreas.  
The additional obs: dist equals zero when not in sight, else the actual dist – likely undesirable.   
OS: 3. TS: 50.

Curriculum: 2.0, Steps: 1.000.000, Training time:

***FullSetUpCertain-1.1.2:***

Default hyperparameters, bounded; MS: 4000.   
VectorObs Only – five additional observations added (Dens, dist) on the CrowdedAreas and dist to target.  
The additional obs: dist equals length of foresight when not in sight, else the actual dist – likely to perform better than the previous attempt. Dens equals zero then not insight, else the actual dens.   
OS: 3. TS: 50.   
Curriculum: 2.0, Steps: 1.000.000, Training time: 10170

***FullSetUpCertain-1.1.3:***

Default hyperparameters, bounded; MS: 4000.   
VectorObs Only – four additional observations added (Dens, dist) on the CrowdedAreas.  
The additional obs: dist equals length of foresight when not in sight, else the actual dist – likely to perform better than the previous attempt. Dens equals zero then not insight, else the actual dens.  
OS: 3. TS: 50.   
Curriculum: 2.0, Steps: 1.000.000, Training time: 9000

**Back on the road:**

FullSetUpCertain-1.2: Default hyperparameters, except: Batch: 512, buffer 256. Bounded; MS: 4000. VectorObs + Grey. OS: 3. TS: 50. Steps: 1.000.000, Training time:

FullSetUpCertain-1.3: Default hyperparameters, except: Batch: 512, buffer 256. Bounded; MS: 4000. VectorObs + Colour. OS: 3. TS: 50. Steps: 1.000.000, Training time:

## Old stuff:

**Parameter-tuning-training-sessions:**

* + **BaseEnvironment-1.2:**
* **DI: 1, TimeScale: 20, TFR: 60, MaxSteps: 0**
* **Training Time: Not recorded**
  + **BaseEnvironment-DI1-TFR-(-1):**
* **DI: 1, TimeScale: 20, TFR: -1, MaxSteps: 0**
* **Training Time:**
  + **BaseEnvironment-DI3:**
* **DI: 3, Timescale: 20, TFR: 60, MaxSteps: 0**
* **Training Time: roughly 3600 seconds**

**A new day, new tuning sessions:**

**Benchmark:**

**Benchmarking-1.0:**

TimeScale: 20  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1082 seconds

**The effect of the number of hidden units:**

**Benchmarking-1.1:**

TimeScale: 20  
DI: 5  
Hidden-units: 512  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1182 seconds

**Overall result:** looks more stable than benchmark, despite the single drop-out – possibly caused by the target being located at one of the difficult regions.

**Try again:** Yes

**Benchmarking-1.2:**

TimeScale: 20  
DI: 5  
Hidden-units: 128  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 947 seconds

**Overall result:** marginally worse than benchmark

**Try again:**  No

**The effect of the timeScale**

**Benchmarking-2.1:**

TimeScale: 10  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1051 seconds

**Overall result:** Decent results, better end performance than the benchmark.

**Try again:** No

**Benchmarking-2.2:**

TimeScale: 40  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 876 seconds

**Overall result:** More stable and better end results than benchmark.

**Try again:** Yes

**Benchmarking-2.3:**

TimeScale: 60  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1023 seconds

**Overall result:** Even better than the previous.

**Try again:** Yes.

**Benchmarking-2.4:**

TimeScale: 80  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1028 seconds

**Overall result:** Rather unstable results, yet better end results than benchmark.

**Try again:** No

**Try 80 again with *collision detection = continuous speculative***

**Benchmarking-2.5:**

TimeScale: 80  
DI: 7  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1100 seconds

**Overall result:** Better than its companion but worse than 40/60 with discrete collision detection.

**Try again:** No

**Benchmarking-2.6:**

TimeScale: 80  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1100 seconds

**Overall result:** Better than its companion but worse than 40/60 with discrete collision detection.

**Try again:** No

**The effect of changing decision interval:**

**Benchmarking-3.1:**

TimeScale: 20  
DI: 3  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1013 seconds

**Overall result:** Reaches same lessons as the benchmark but performs worse in the beginning yet more stable at the end.

**Try again:** Yes

**Benchmarking-3.2:**

TimeScale: 20  
DI: 7  
Hidden-units: 256  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time: 1065 seconds

**Overall result:** Appears more unstable than the benchmark.

**Try again:**  No

**The effect of changing batch/buffer size:**

**Benchmarking-4.1:**

TimeScale: 20  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 1024/5120  
LR: 3.0e^-4  
Training time: 1124 seconds

**Overall result:** Achieves better end result than benchmark, despite two dropouts (possibly because of difficult-area-syndrome).

**Try again:** Yes

**Benchmarking-4.2:**

TimeScale: 20  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 4096/20480  
LR: 3.0e^-4  
Training time: 1097 seconds

**Overall result:** Far worse than benchmark, no gets passed lesson 0.

**Try again:** No

**Candidates for final:**

**Candidate-1.0:**

TimeScale: 60  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 1024/5120  
LR: 3.0e^-4  
Training time (Static): 1717 seconds (With three runs)  
Training time (Clouds): 1840 seconds (With three runs)  
Training time (Pedestrians): 1745 seconds (With three runs)

Training time (full set-up, 1 mil. Steps): 5600 seconds

**Overall result:** Really promising, reaches the final lesson and somewhat stable training.

**Candidate-1.1:**

TimeScale: 40  
DI: 5  
Hidden-units: 256  
Batch/Buffer: 1024/5120  
LR: 3.0e^-4  
Training time (Static): 1020 seconds  
Training time (Clouds): X seconds (With three runs)  
Training time (Pedestrians): X seconds (With three runs)

**Overall result:** Too much variability.

**Candidate-1.2:**

TimeScale: 60  
DI: 5  
Hidden-units: 512  
Batch/Buffer: 1024/5120  
LR: 3.0e^-4  
Training time (Static): 1970 seconds (With three runs)  
Training time (Clouds): 1850 seconds (With three runs)  
Training time (Pedestrians): 1945 seconds (With three runs)

**Overall result:** Overall decent results, yet 1.0 seems better – Higher average value-estimate-path and less variability in lessons – Cum. Reward and episode length appears to be similar.

**Candidate-1.3:**

TimeScale: 60  
DI: 5  
Hidden-units: 512  
Batch/Buffer: 2048/10240  
LR: 3.0e^-4  
Training time (Static): 1930 seconds (With three runs)  
Training time (Clouds): X seconds (With three runs)  
Training time (Pedestrians): 1900 seconds (With three runs)

**Overall result:** Not impressive, never gets to the final lesson. ***(Could perhaps benefit from lower beta value, as slower decrease in entropy is witnessed)***

**Candidate-1.4:**

TimeScale: 40  
DI: 5  
Hidden-units: 512  
Batch/Buffer: 1024/5120  
LR: 3.0e^-4  
Training time (Static): X seconds  
Training time (Clouds): X seconds (With three runs)  
Training time (Pedestrians): X seconds (With three runs)

**Overall result:** Too much variability again.

**Other interesting parameters:**

* Time horizon

**Runs:**

*Dissertation-3.2:* A full run with the target spawning randomly at six different places – sensor clouds included but no moving pedestrians.

*RandomPlacingTarget-1.0:* A full run with the target randomly spawning anywhere within the boundaries of the environment. Sensor clouds included but no moving pedestrians.

*FullSetupTest-1.4:* First full run with the full implementation.

*Profiling-1.8:* Run with trails being drawn without time delay. (Not a full run because of the expensiveness of drawing trails – on going challenge to be solved)

*Profiling-1.9:* Run with trails being drawn with a time delay of 0.5 seconds. (Not a full run because of the expensiveness of drawing trails – on going challenge to be solved)

*CurriculumNavigation-1.8:* First run where the curriculum setup worked (Single area).

*FInalMultiAreaSetUp-1.0:* First run where the curriculum setup (multi area) worked as intended. However, unfortunately the sensor clouds did not change position in the final lesson, which is suspected to have affected the results.

*FInalMultiAreaSetUp-1.1:* First run where the curriculum setup (multi area) worked as intended.