

Week -1 (18.10 – 24.10)

19/10

- Start-up, watched presentation, setup basic structure for project work folder.

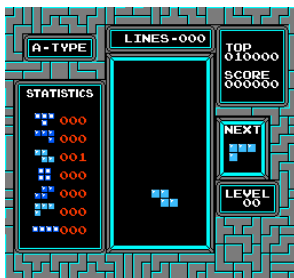
Week 0 (25.10 – 31.10)

- Busy due to assignments and obligatory coursework in other courses.

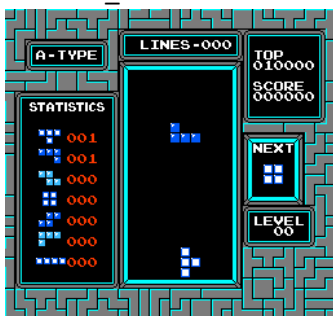
Week 1 (01.11 – 07.11)

Thursday 04/11

- Chose task Reinforcement Learning.
- Need to decide on environment. Dabbling in the idea of creating own, but this could take a while to implement so settling for a pre-made one for the time being.
 - <https://github.com/Kautenja/gym-tetris>
- Trying to prepare dependencies etc., encountering issues with Visual C++.
- Got the environment running.



- Experimenting with the environment to figure out the variables. Using the “SIMPLE_MOVEMENT” action space since it has all the inputs we require for Tetris.



- For future reference:

| | |
|-----------|--------------------------------------|
| 0 – NOOP | No Operation, agent idles |
| 1 – A | Agent rotates piece clockwise |
| 2 – B | Agent rotates piece counterclockwise |
| 3 – right | Agent moves piece to the right |
| 4 – left | Agent moves piece to the left |
| 5 – down | Agent soft drops piece |

Friday 05/11

- Decided on trying to implement the AI using DQN and referenced the 8th exercise. Quickly ran into a wall and decided to start looking into how to save and load trained models before doing anything else, as this will probably prove to be rather important later, since Tetris is significantly more complex than the CartPole environment. I therefore expect the training to take quite some time.
- Noticed some flawed logic in the CartPole DQN model, specifically with the exploration rate, where the agent would never explore anything. This leads to the agent only being able to evolve if it gets “lucky” in the first few episodes.
- Implemented a fix that lets the agent learn more consistently, although it does not always retain what it learns.

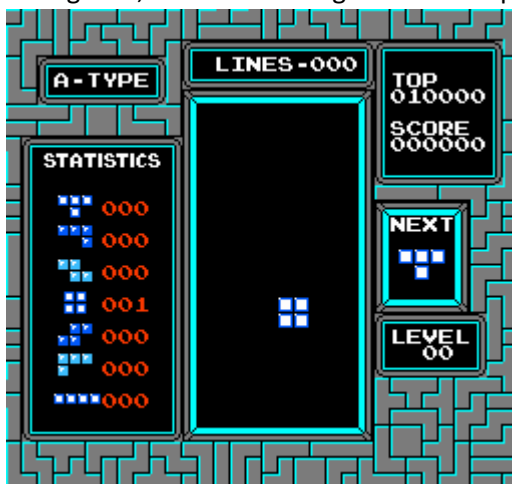
Week 2 (08.11 – 14.11)

Wednesday 10/11

- Noticed that the previous fix seems to have broken the model somewhat, in the sense that it is no longer able to continue from where it was after saving and loading.
- Researched some others’ implementations of Tetris AI, seems that most people create their own environments.
- Attempts at fitting the model to the chosen environment, encountering difficulties as there isn’t really any documentation, which makes it hard to figure out what exactly each component is.

Friday 12/11

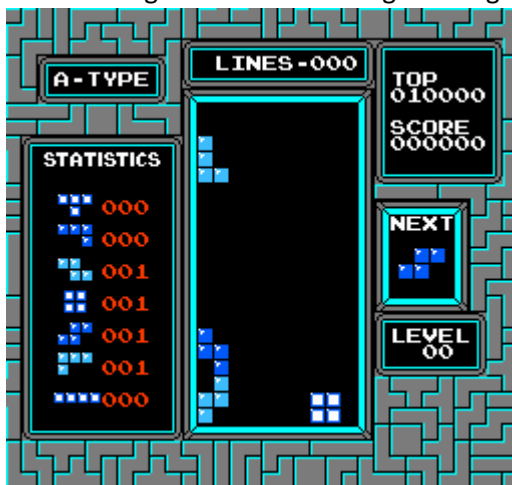
- Resumed attempts at trying to fit model. Trying to figure out what exactly the state returns etc.
- Attempting to use the “info” that the environment returns instead, as there is some documentation on what these are. A simple solution lets the agent play with some degree of intelligence, but I could not get it to learn properly.



Week 3 (15.11 – 21.11)

Tuesday 16/11

- Trying to fit the model again. Finally realized that the 240 x 256 x 3 state that the environment returns is each individual pixel of the 240 x 256 window, as well as the RGB color values.
- Had a meeting with Jonathan, who recommended trying a different environment since the one I had chosen wasn't the greatest to work with. Introduced an environment developed by last year's students of TDAT3025, <https://github.com/JLMadsen/TetrisAI>, and suggested using this one and focusing on some interesting way of implementation instead of the developing of an environment in itself.
- Decided to play around with the [Kautenja](#) environment for a little longer just to see where I'm at before I move on. Left the agent to play for a large amount of time and recorded its progress, but when I came back it had crashed on episode 24. After reviewing the footage, I found that the crash was caused by the agents first line clear, which is very unfortunate as this was a chance to see if the agent would evolve after finally finding a way to increase its score.
- Experimented with letting the agent get rewarded simply by staying alive, but this seemed to make the agent rather hyper-active. It's likely that the agent understood this reward as "press as many buttons as possible" instead of "build as low as possible to stay alive", since the reward gained from mashing outweighed the penalty from stacking high.



- Concluded that my general understanding of Deep and Reinforcement Learning is somewhat lacking, so I'm going to spend some time getting a grip on these topics before moving on to the new environment.

Week 4 (22.11 – 28.11)

Assignment turn-in 26.11!