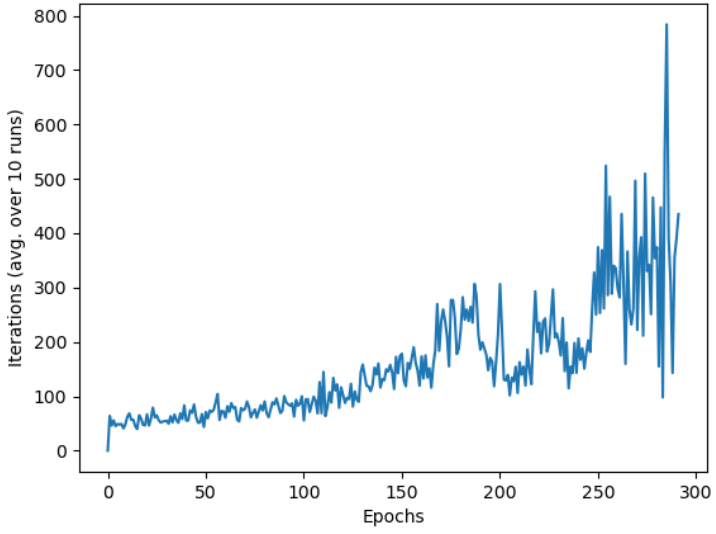
ATAI-720-2020 Assignment 1

Summary of results and general remarks

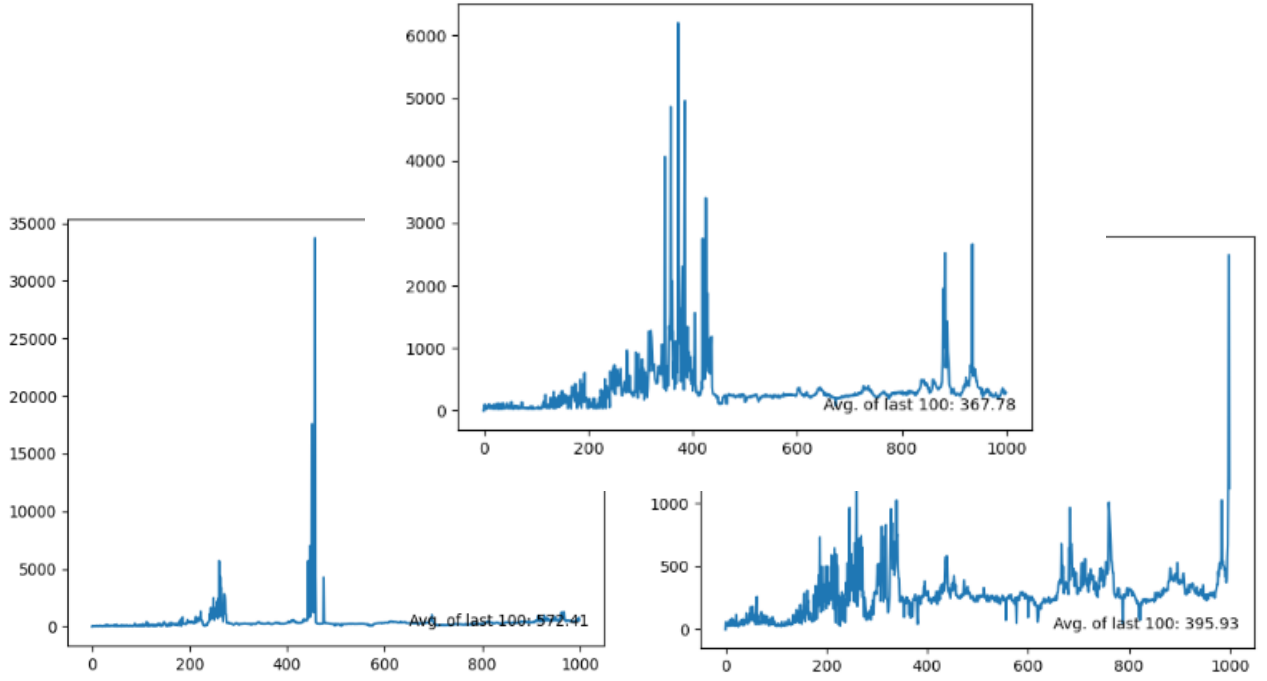
Plain Vanilla

- All tests seemed successful, used as a “base-line” for further tests.

Average



Single Runs



Small noise on variables

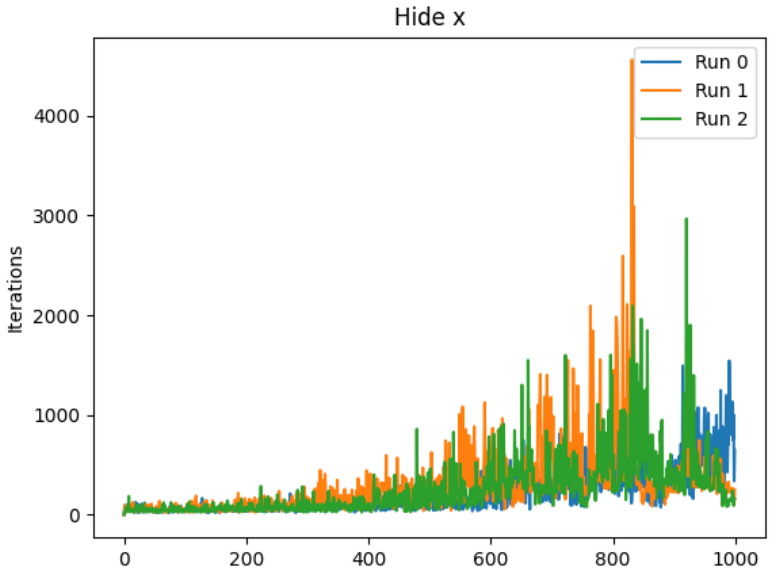
- Generally has impact on performance depending on the noise level.
- The higher the noise the worse the performance
- Noise on action values (noise on the +-10 N) has little impact

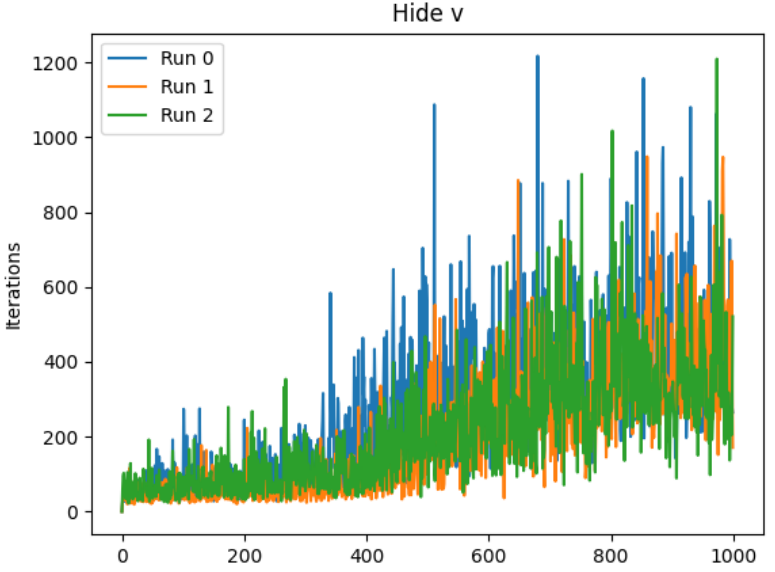
Discretization

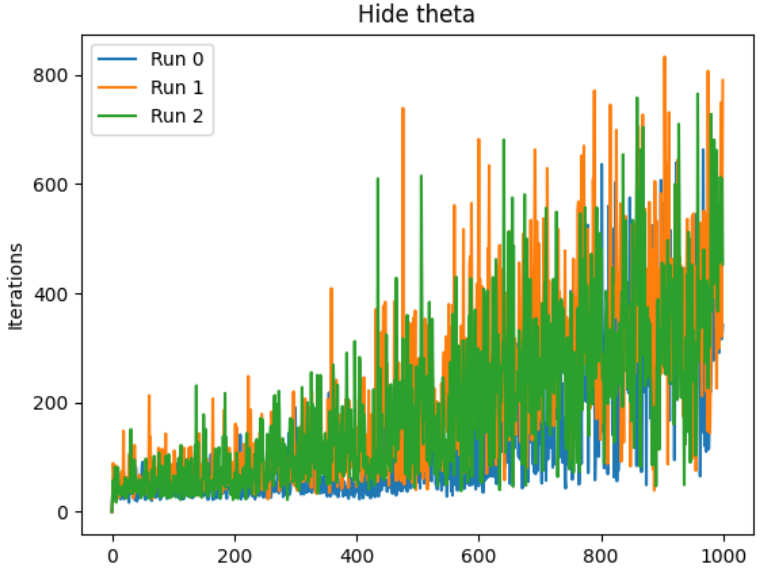
- Discretization in time: lower time rate reduces performance (as expected), higher time rate (smaller dt) increases performance.
- Rounding to next .1 in observation showed little impact
- One of you tested discretization “by importance” meaning higher resolution in the center (+-6°) and low resolution everywhere else → The learner learned quicker.

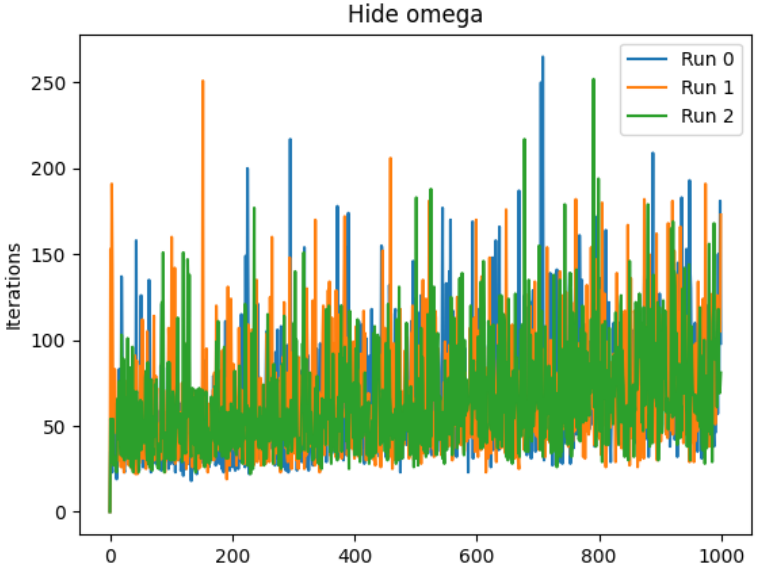
Hidden variables

- Consensus, that
  - Hiding x doesn’t change the success of the learner
  - Hiding theta/ v reduces learning
  - Hiding omega reduces learning the most


Hide x
Run 0
Run 1
Run 2
Iterations
0
1000
2000
3000
4000
0
200
400
600
800
1000

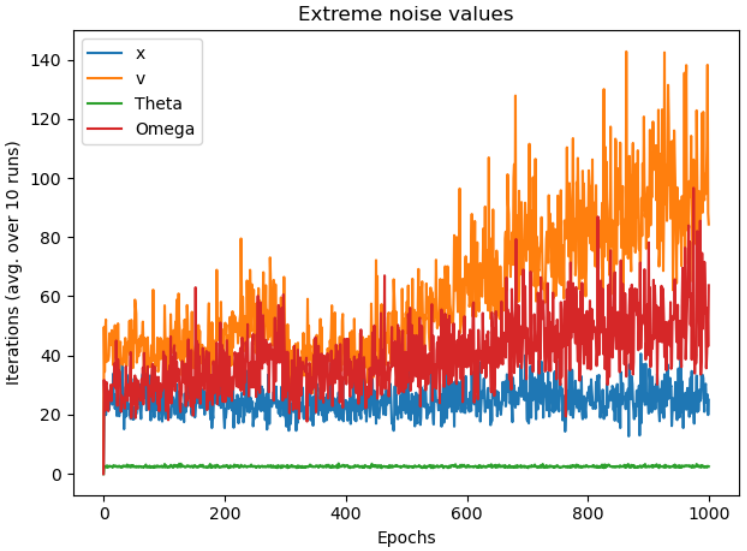

Hide v
Run 0
Run 1
Run 2
Iterations
0
200
400
600
800
1000
1200
0
200
400
600
800
1000


Hide theta
Run 0
Run 1
Run 2
Iterations
0
200
400
600
800
0
200
400
600
800
1000

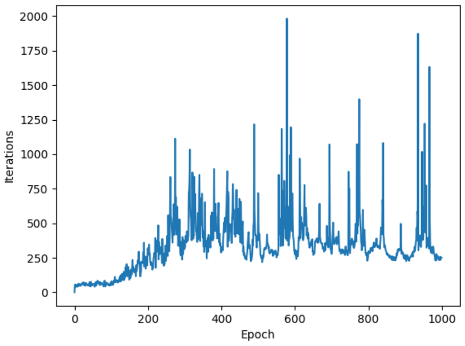

Hide omega
Run 0
Run 1
Run 2
Iterations
0
50
100
150
200
250
0
200
400
600
800
1000

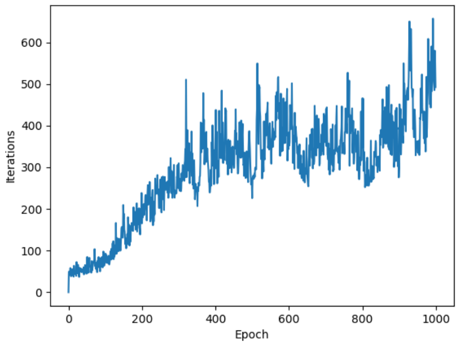
Strong noise on single variable

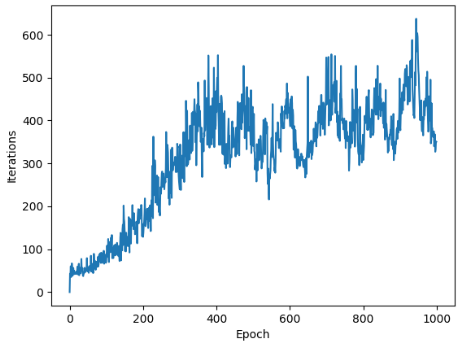
- Has high impact on learning
- Interestingly has higher impact on learning, than hiding a variable – Quick Question: What does this mean for AI – for autonomy – and for “fairness”?

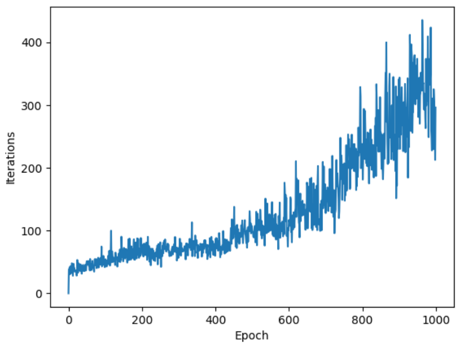


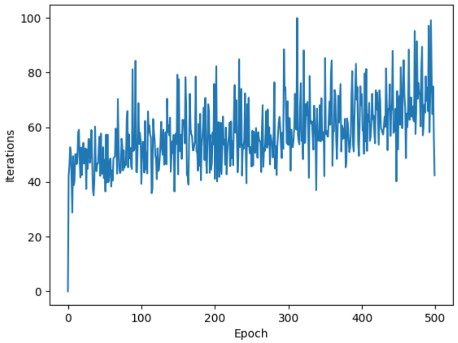
- Hidden vs. Noisy: x, v, theta, omega (in that order)

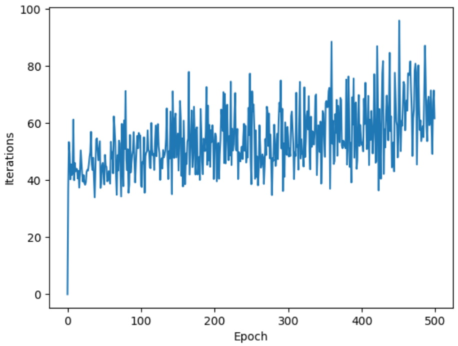


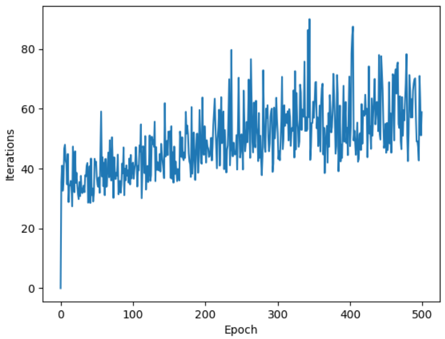


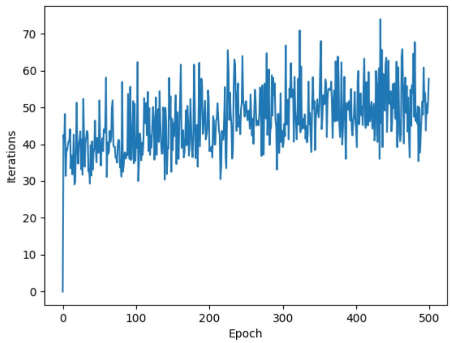






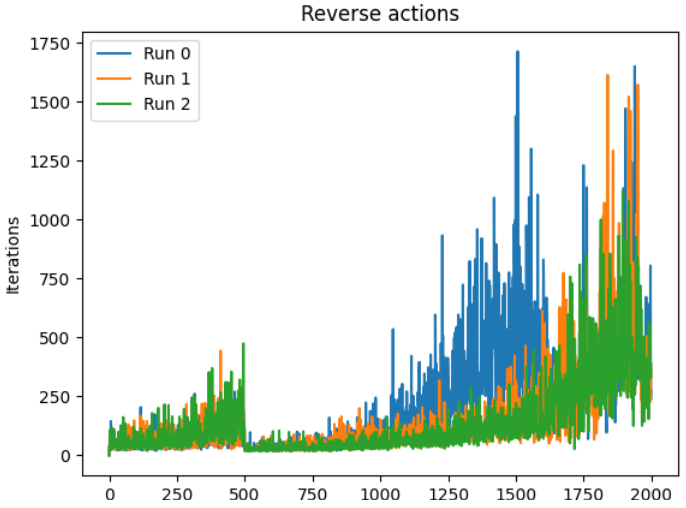


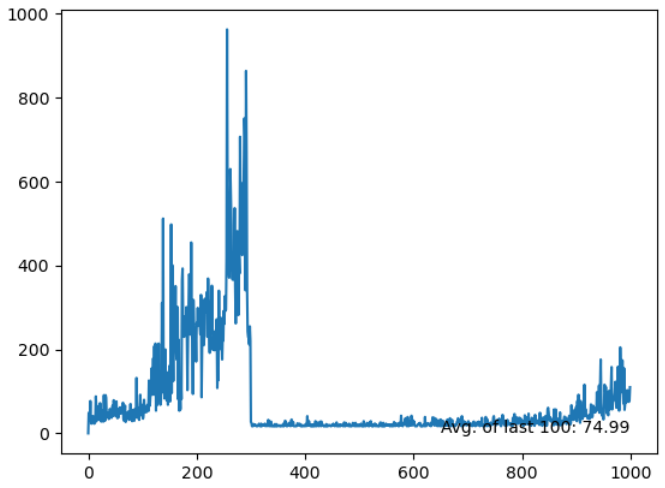


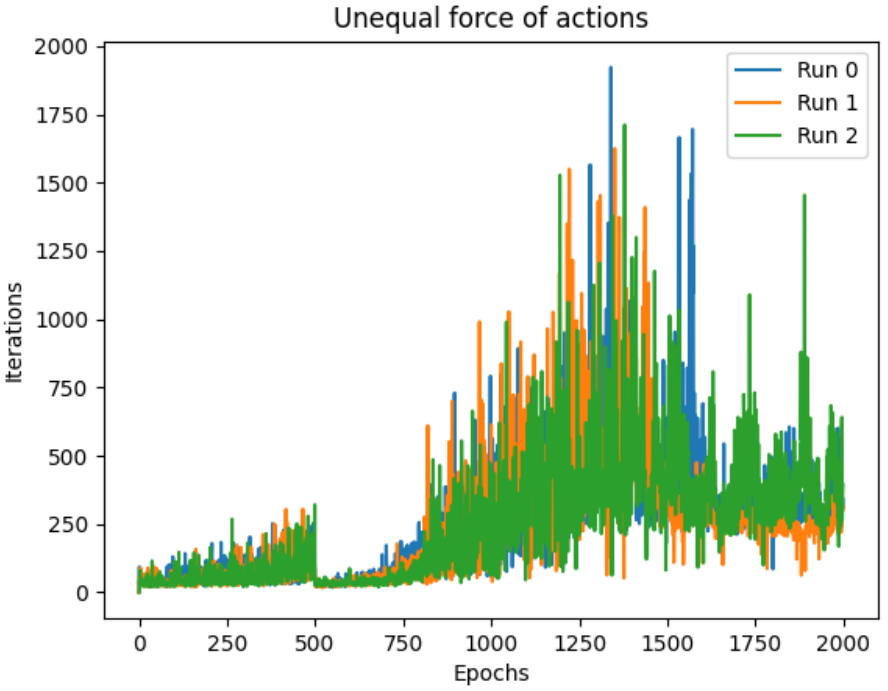


Changes to the cart-pole

- What has been done:
  - Inversion after certain number of epochs → see next slide
  - Increasing of gravity → little impact
  - Increasing of pole / tip of pole mass → little impact
  - “Uneven” forces (F = [20, -5]) → see next slide


Reverse actions
Run 0
Run 1
Run 2
Iterations
0
250
500
750
1000
1250
1500
1750
2000


Avg. of last 100: 74.99
0
200
400
600
800
1000


Unequal force of actions
Run 0
Run 1
Run 2
Iterations
Epochs
0
250
500
750
1000
1250
1500
1750
2000

Discussion

- What most of you concluded on:
  - RL is similar to human learning, but is lacking important features (No pure trial and error in human learning)
  - The AC is not good at coping with novelty
  - Changes in observation space cannot be coped with
  - Novelty can cause the learner to have to not only learn from scratch but even further back (has to “unlearn” previous knowledge before relearning)

General remarks

- Plots really help to understand what you are talking about
- Think about what information can be useful, research it if necessary (e.g. learning rate, performance) and don’t take single test results as a result for the whole thing, try it out a couple of times, average data, calculate std-devs, make boxplots, whatever helps you to visualize the results.
- The assignments usually don’t have a correct answer, their purpose is to make you think about current AI and possibilities to advance the state of the art.
- Your own opinion is not only allowed, but really requested in those assignments.
- You don’t have to write up *everything* you have done. Try to keep it short (e.g. write down the parameters and the results – with some explanation if possible) No need for long introductions or detailed description of a parameter change, rather write detailed about the results and what they suggest). A short description how you ran it, which machine you used etc. is always useful.
- If code snippets help to explain what you have done include them, otherwise leave them out