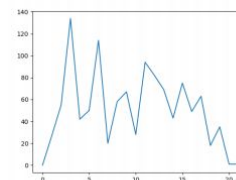
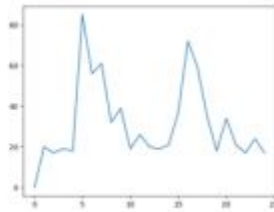
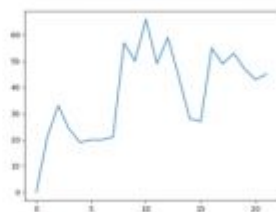
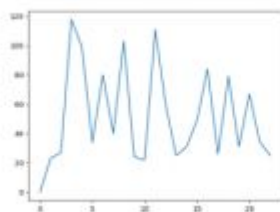
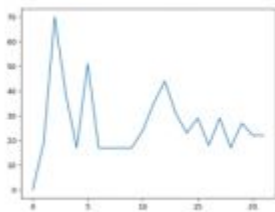


# ATAI-720-2020 Assignment 2

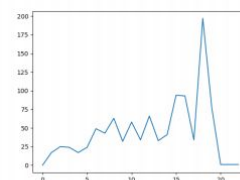
Summary of results and general remarks

# Plain Vanilla

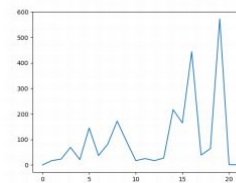
- Agreement, that previous knowledge is being brought into the game.
- Assignment 1 gave you information about failure conditions etc. -> not a blank sleeve as the actor-critic was.
- Generally most of you best in condition 3, worst in condition 1.



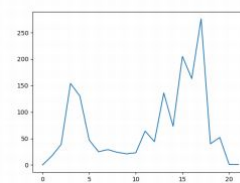
1. SYMBOLIC  
Hi: 134  
Avg: 59.105  
Med: 55



2. COLOR  
Hi: 198  
Avg: 55.632  
Med: 42

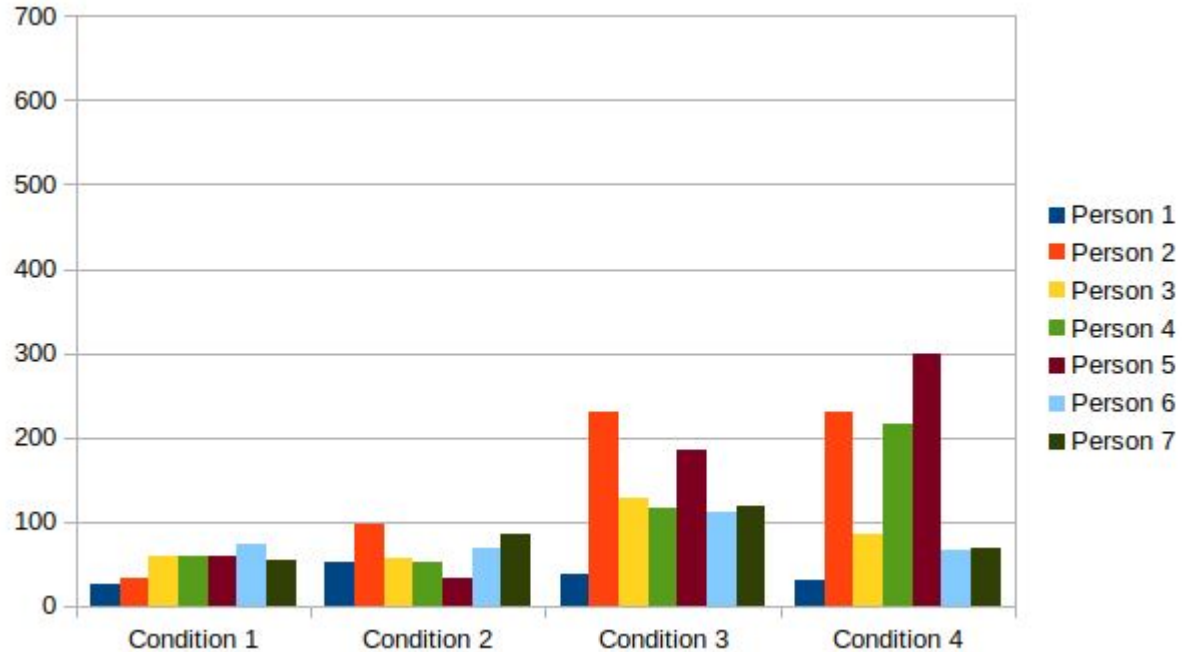


3. LINES  
Hi: 582  
Avg: 127.842  
Med: 51



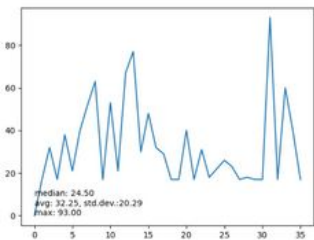
4. ANIMATION  
Hi: 272  
Avg: 86.263  
Med: 48

# Plain Vanilla

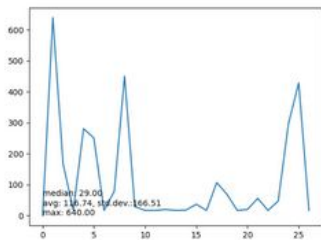


# Force Inversion

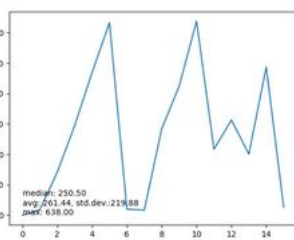
- Agreement, that knowing of the inversion gives advantage over the actor-critic.
- Human learner is not only a learner, but also “exists outside the task”.
- Attention is important.
- But too fast to do conscious decisions.



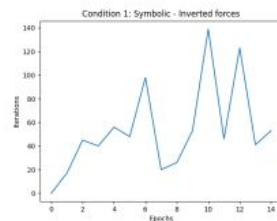
1. Plain



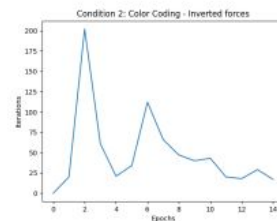
2. Colors



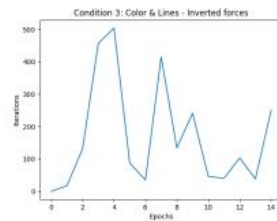
3. Colors line



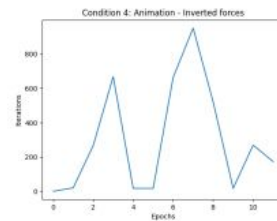
(a) Condition 1: Symbolic



(b) Condition 2: Color Coding



(c) Condition 3: Color & Lines



(d) Condition 4: Animation

# Hidden theta

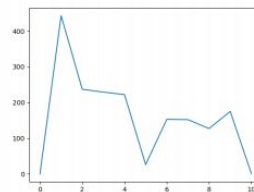
Most of you agree, that this makes the task much harder.

Interestingly this is different to the AC (which primarily uses omega).

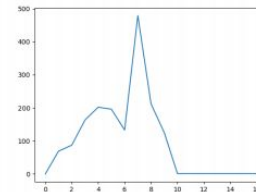
Most of you primarily looked at theta, not at omega.

Might this be a local minimum?

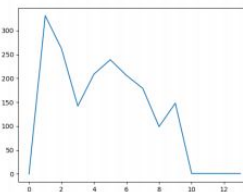
Condition	Max Score	Average	Std. Deviation	Median
Plain	42	24	11.49	21
Colors	112	41.05	29.22	31
Color lines	138	49.44	31.55	47.5
Animation	295	90.8	89.81	74



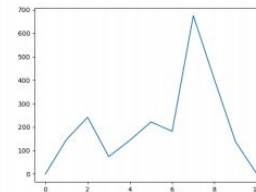
1. SYMBOLIC  
Hi: 443  
Avg: 196  
Med: 175



2. COLOR  
Hi: 424  
Avg: 189.222  
Med: 163



3. LINES  
Hi: 331  
Avg: 201.667  
Med: 206



4. ANIMATION  
Hi: 675  
Avg: 247.222  
Med: 182

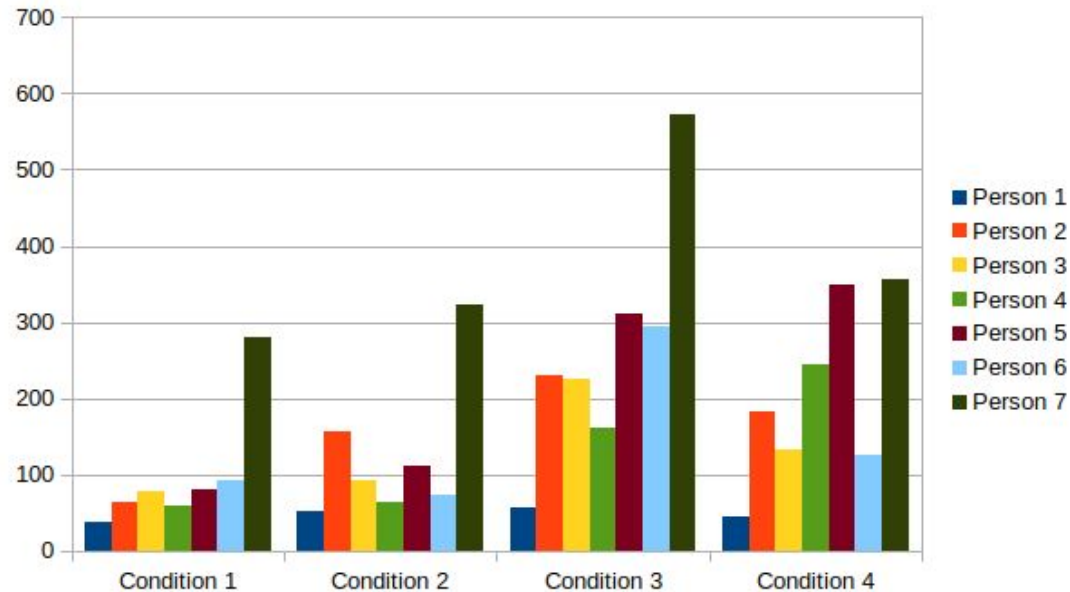
# Your modifications

- Changes were done to the environment. Some of you chose them as in assignment 1.
- Including hidden vs random:  
Humans can distinguish random variables and don't use them for decision making
- Or uneven forces:  
Makes the task much harder
- Or synchronous environment:  
Makes the task much easier. No failure at all sometimes. Almost indefinitely long runs.

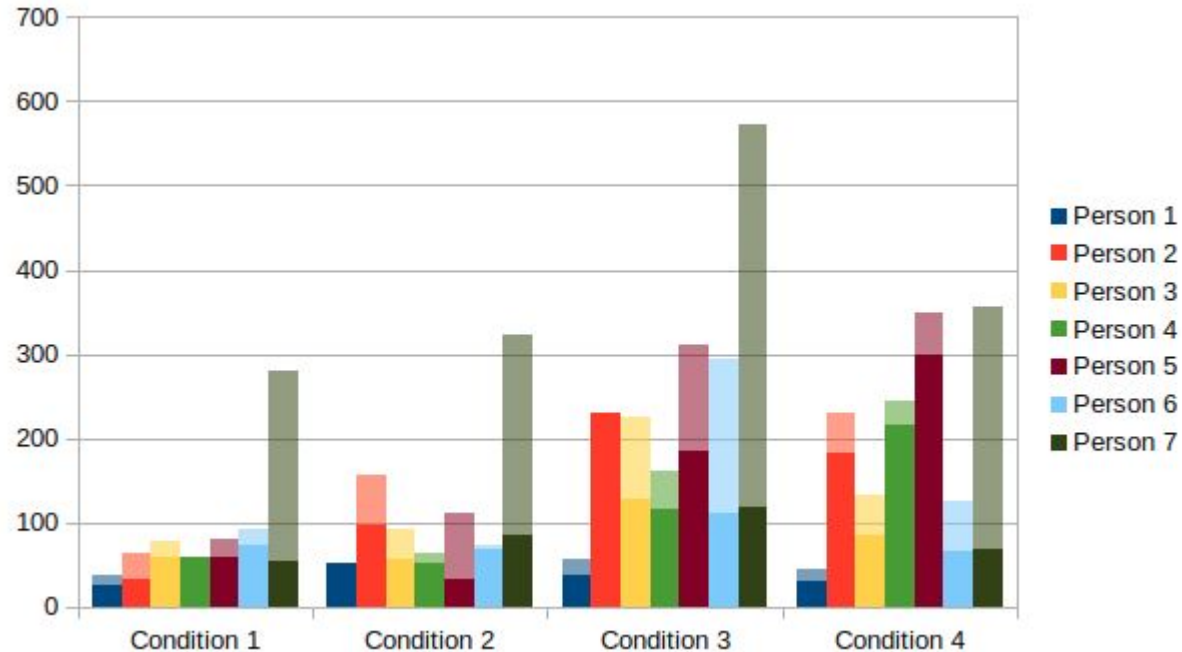
# Vanilla 2

All of you performed much better in the last task.

Cumulative learning helped you to perform better in the task.



# Comparison Vanilla 1 and 2





# Discussion - Your thoughts

- Boredom makes the task annoying at times.
- Outside stimuli make you lose focus and lead to failure.
- Cumulative learning is a huge advantage over simple reinforcement learning.
- Attention is extremely important.
- Knowledge can be transferred between different tasks.
- If the task is very constraint and without changes a RL will outperform a human.

In Assignment 1 many of you concluded, that RL is similar to human learning. Would you stick with that?

# General Remarks

- Again plots really are useful. If you talk about that, for example, after the inversion the performance dropped for a short period it would be nice to see this result in your data.
- Try to provide all data (in the form of plots, tables etc.) that you use in your text.
- Read the instructions carefully! Do the parts that are described in detail the way they are described (e.g. don't add to it or change it - do this in the section where you are supposed to try things that you are curious about).
- Think about “high-level” conclusions. What impact have certain properties on other parts of the learner (e.g. human)? Ask yourself: What does this mean for learning in general?
- Don't only list your results. Discuss them as well.