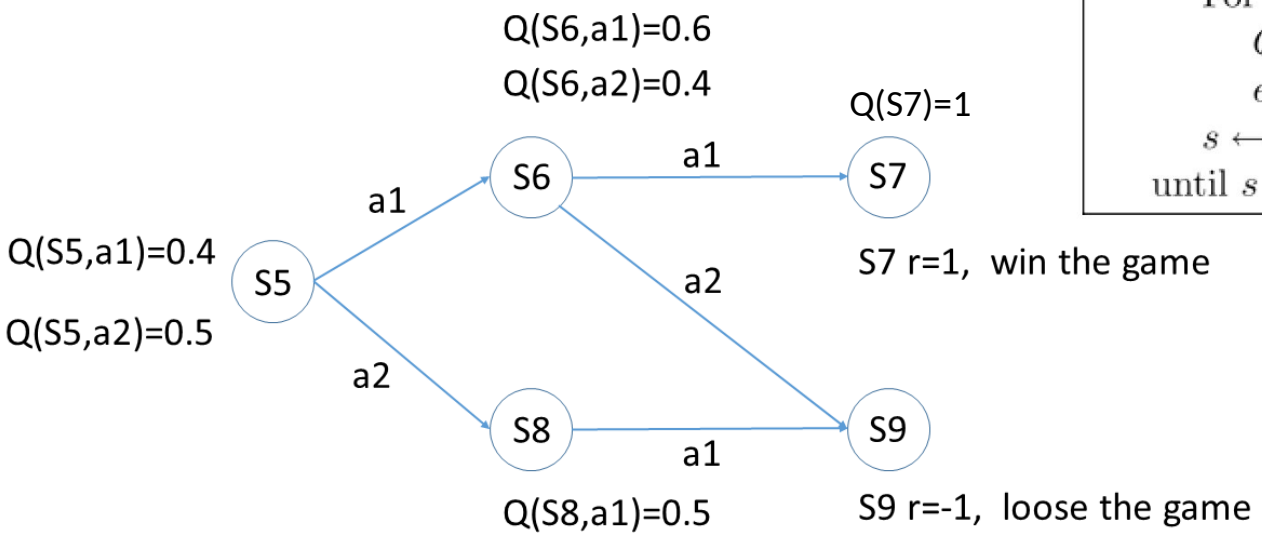
Tricking your cerebellum

Exercise 3 – Feedback

Exam question RL:

•Update the tree (6p)
Take the below described sarsa(λ) algorithm to determine step-wise the updates to Q-values of state s5 for 2 episodes.
Take in each state the action with the highest Q-value.
Set parameters λ=0.9,γ=0.9,α=0.9
What are the characteristics of the learning process, when we set λ=0 ?



SARSA(λ)

Initialize $Q(s,a)$ arbitrarily and $e(s,a)=0$, for all s,a
Repeat (for each episode):
  Initialize s, a
  Repeat (for each step of episode):
    Take action a, observe r, s'
    Choose a' from s' using policy derived from Q (e.g., ε-greedy)
    $\delta \leftarrow r + \gamma Q(s',a') - Q(s,a)$
    $e(s,a) \leftarrow e(s,a) + 1$
    For all s,a:
      $Q(s,a) \leftarrow Q(s,a) + \alpha\delta e(s,a)$
      $e(s,a) \leftarrow \gamma\lambda e(s,a)$
    $s \leftarrow s'$; $a \leftarrow a'$
  until s is terminal

Your solutions

Some general remarks

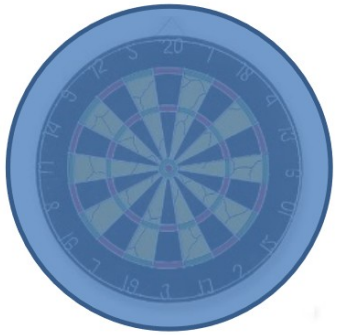
- Please use the correct file names:
  - Bavarian_Game_2026_Ex**<no>_<lastname1>_<lastname2><_code>**
- I need your full names and matriculation number once
- Please include only relevant code in the main document or mark your changes clearly
- Make sure to have >=2 unbiased subjects
- Mark excluded trails in your plots

Motivation feedback

Motor Learning

- Internal Models → error-based learning
  - task-specific prediction error (e.g. direction & amplitude) is used to update an internal model of how this result came about
- Reinforcement Learning
  - an external or internal reward signal is used to improve

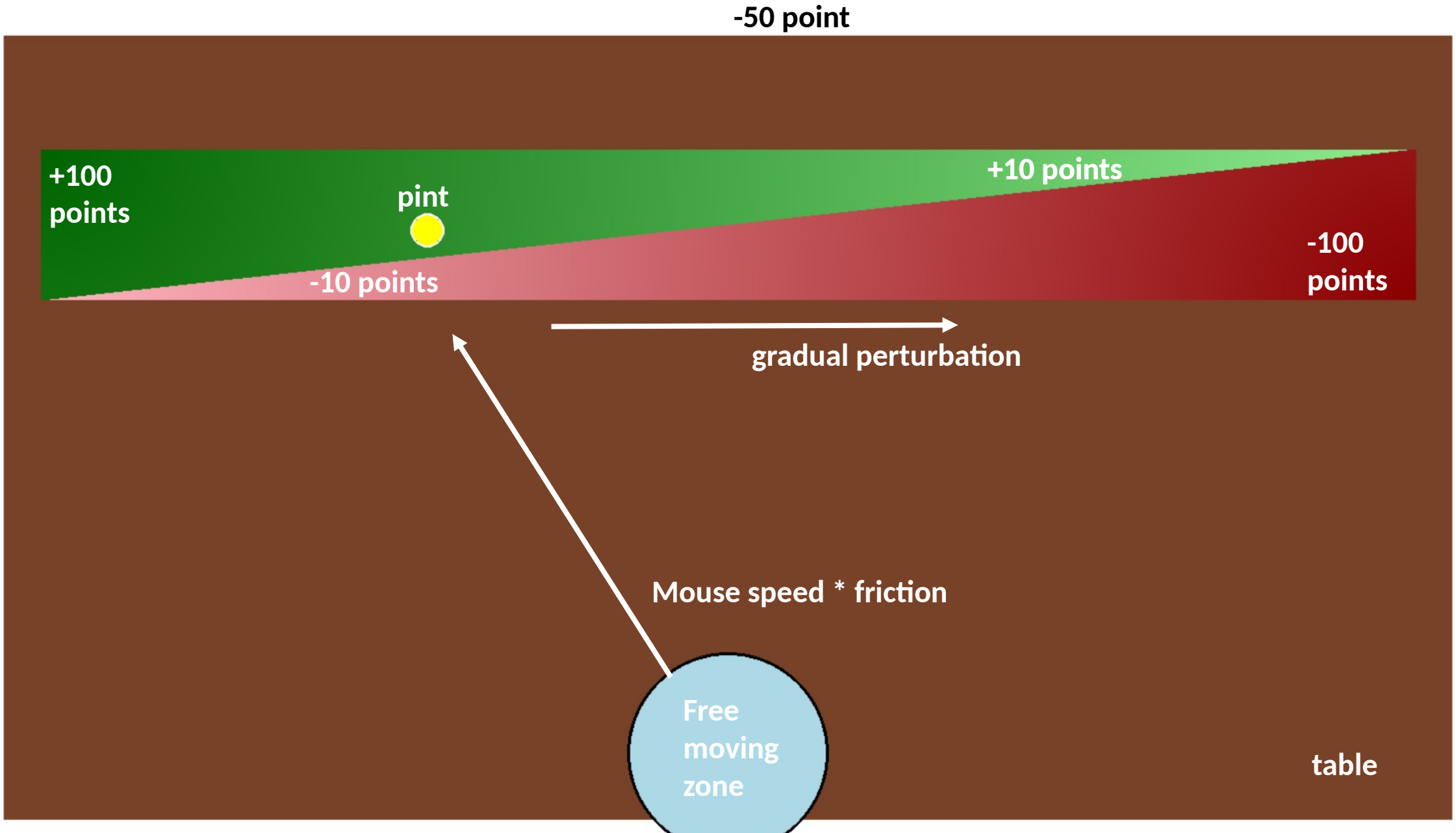
amplitude and direction of error available



Bulls eye: Yes or No

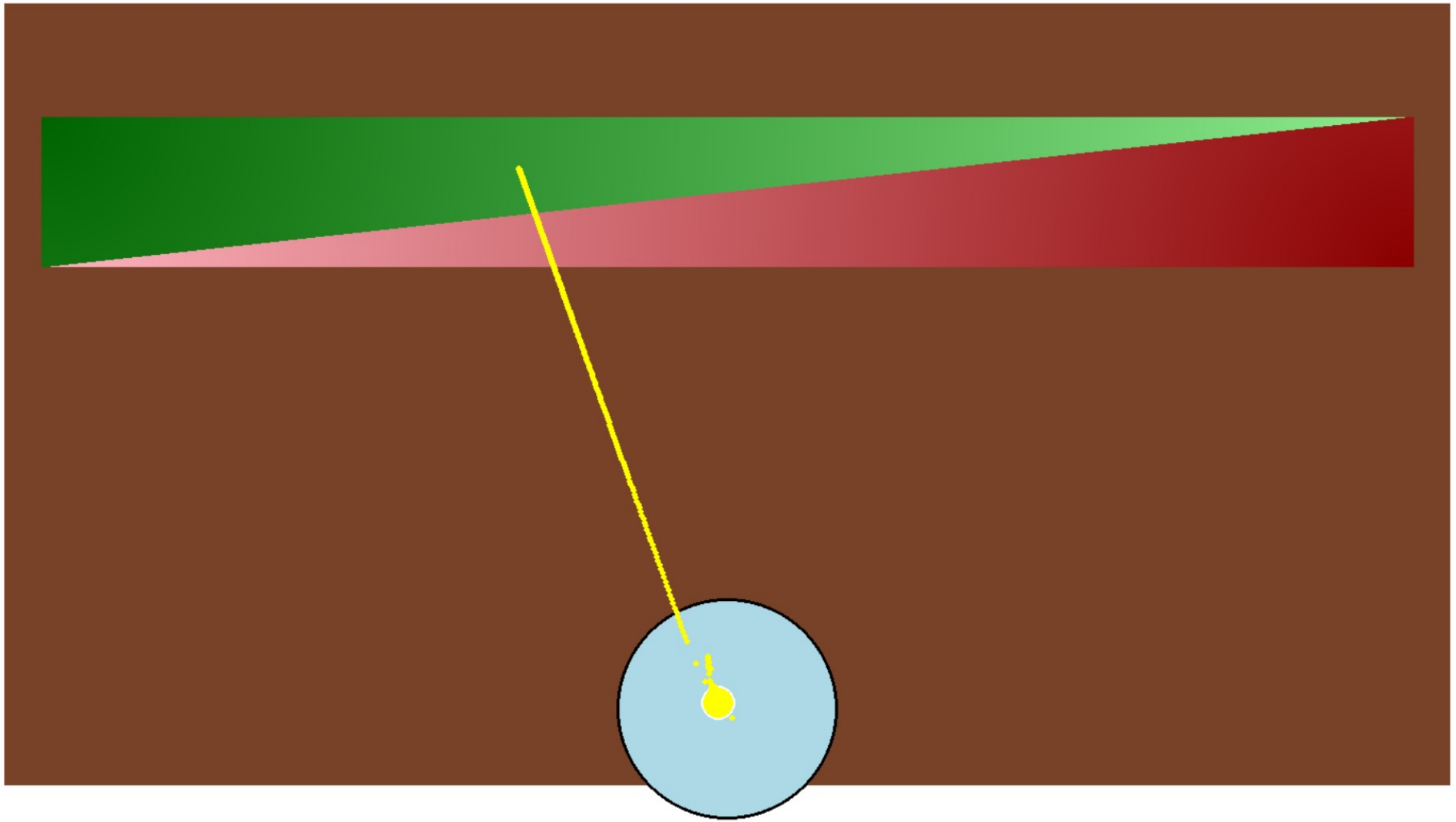
The Bavarian Game: Maßkrugschieben

The Bavarian Game



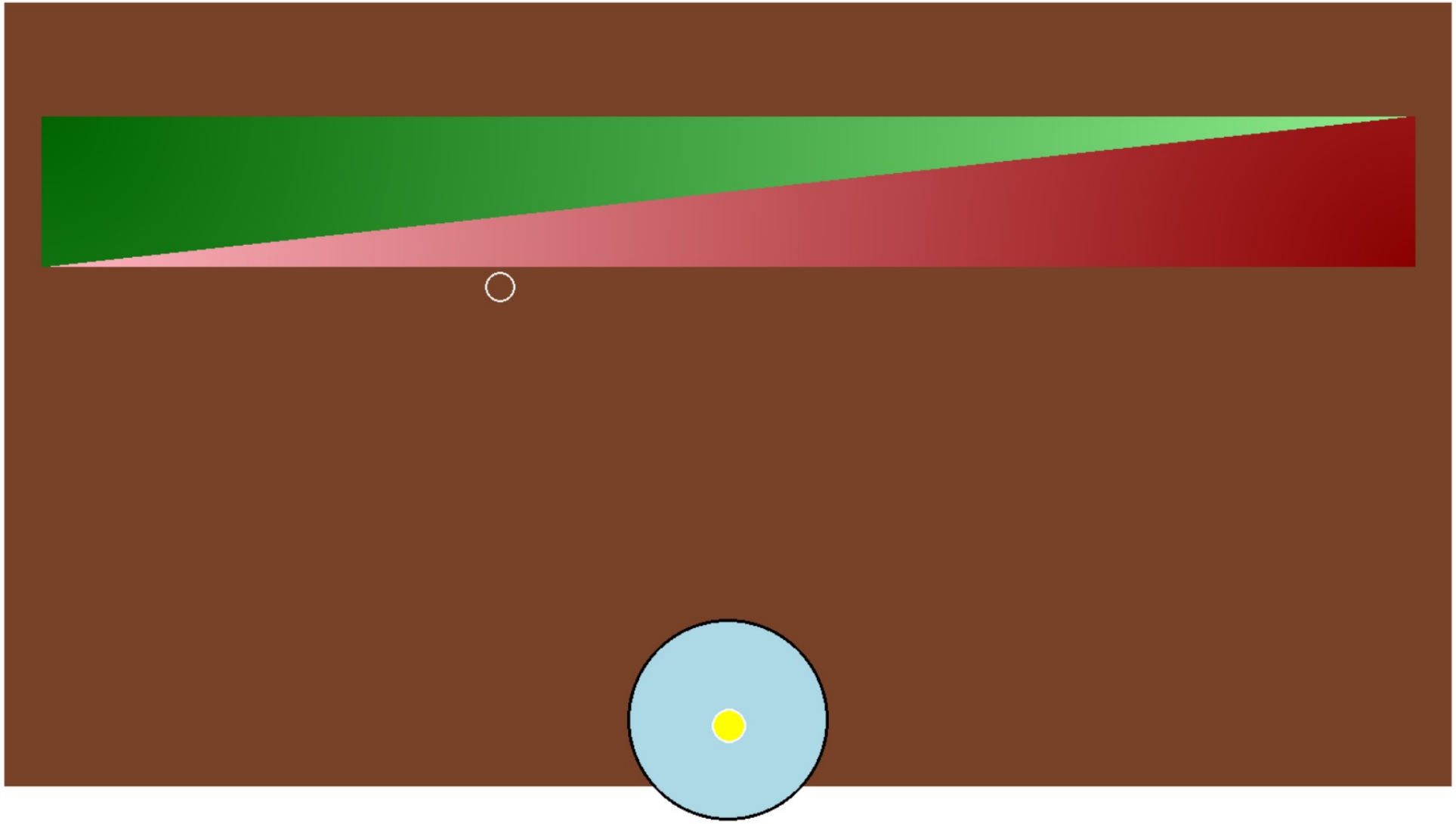
feedback_type=‘trajectory‘

Draw trajectory of the **last** attempt



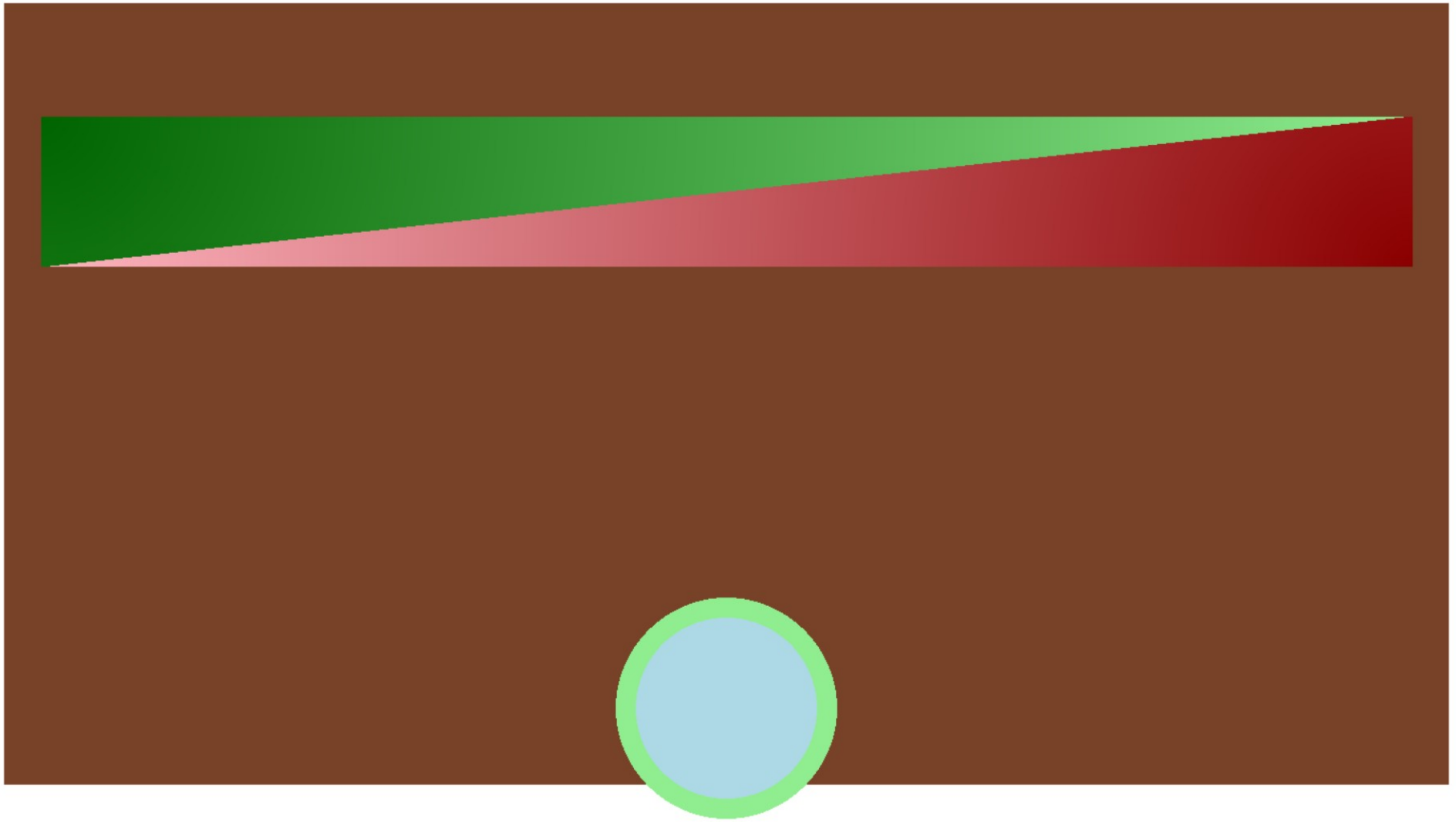
feedback_type=‘endpos‘

Draw pint_end_position of the **last** attempt



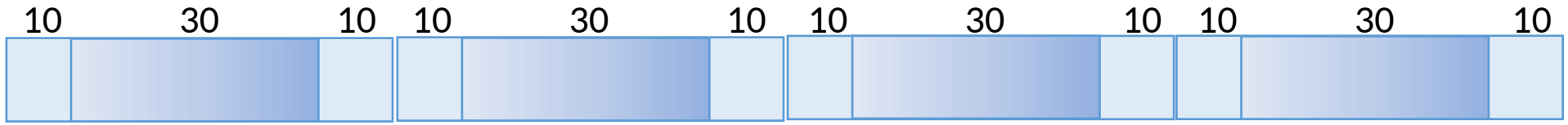
feedback_type=‘rl‘

Change color of starting_postion edge:
RED for miss
GREEN for hit



Design your experiment!

No perturbation
gradual perturbation
.2 in 10 steps, increase after 3 attempts



feedback_type= ‘None‘

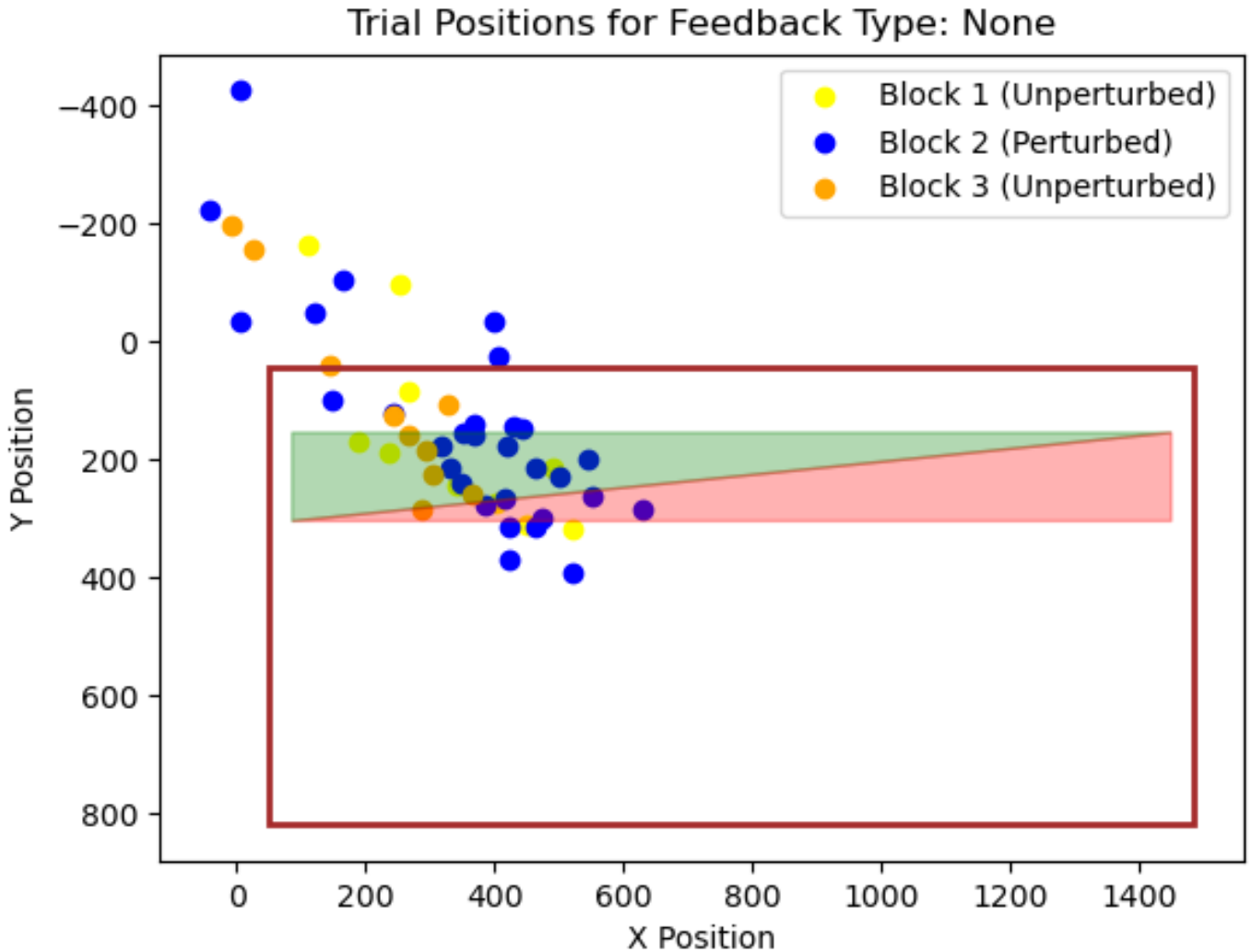
feedback_type= ‘trajectory‘

feedback_type= ‘endpos‘

feedback_type= ‘rl‘

Exercise 3

- TASK 1: Implementation of feedback_types
  - Implement the **3 different feedback_types**: ‘trajectory‘, ‘endpos‘, ‘rl‘
  - **Design** your own experiment to test feedback_types
- TASK 2: Analysis of feedback on **unbiased** subjects (>=2)
  - **Record** the subject performing your own experiment
  - **Visualize** the effect of different feedback on subject‘s performance
- TASK 3: Discussion of your results
  - What’s the effect on subject’s performance of each feedback_type
  - Under which feedback_type was your subject able to adapt the best?
- Bonus TASK 4: Try out your ideas, just state what you did and why it was interesting to you


Trial Positions for Feedback Type: None
Block 1 (Unperturbed)
Block 2 (Perturbed)
Block 3 (Unperturbed)
−400
−200
0
200
400
600
800
Y Position
0
200
400
600
800
1000
1200
1400
X Position