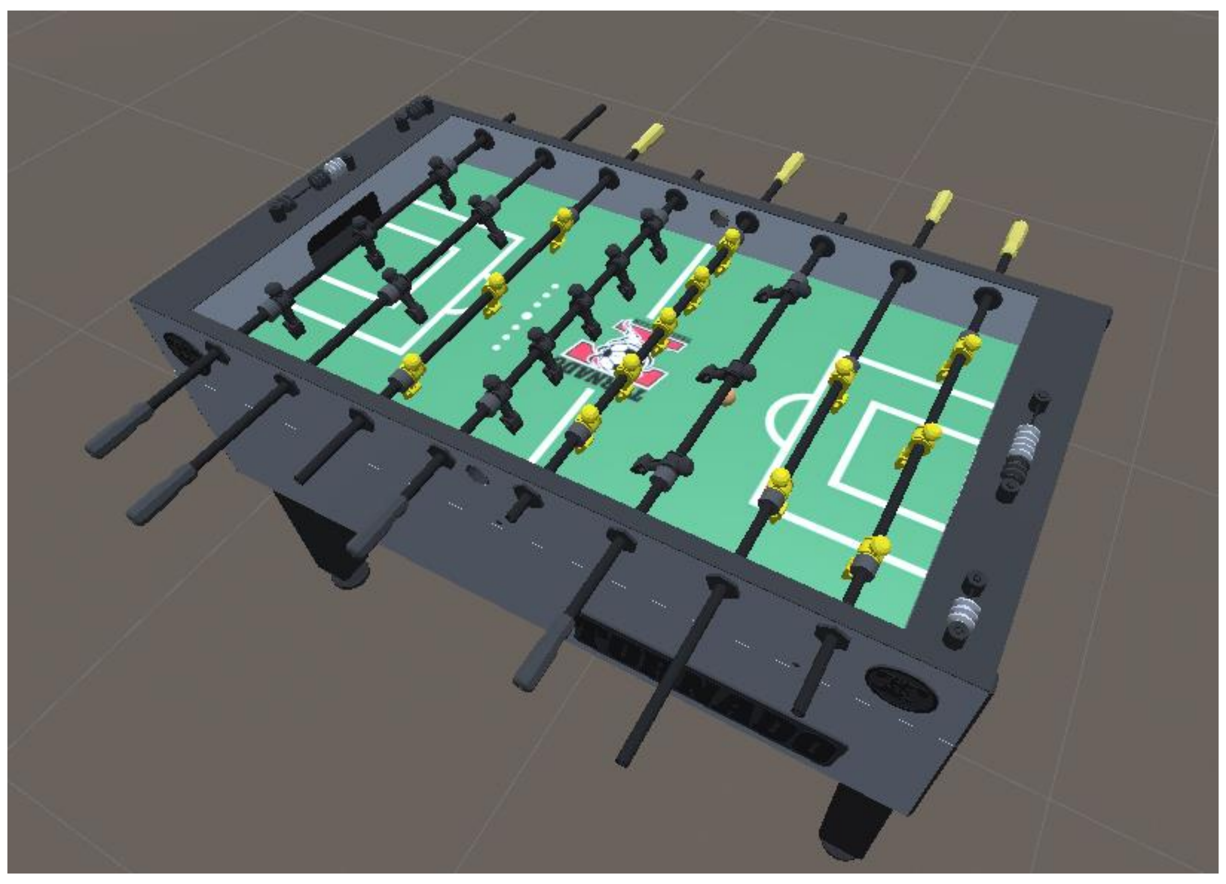
Reinforcement Learning of Action Sequences in Table Football

Author: Alexandros Kyrloglou. Promotor: Wannes Meert. Guidance: Jens Bürger, Wen-Chi Yang.

Introduction:

- Define Action Sequences in Table Football.
- Passing between controllable players.
- Exploration of hybrid[1] Action space and its effect on the performance of the Agent.



Problem Statement:

Action Type	Rota-tion	Trans-lation	Rot. Speed	Tran. Speed	On/Off Move
Discrete all	D	D	constant	constant	-
DCCS	D	D	C	C	-
DCCM	C	C	-	-	D
Continuous all	C	C	-	-	-

D = Discrete, C = Continuous

Types of passes:

- Ball control (tick-tacking)
- Forward Pass 5 to 3 player:
  - vs Static opponent: Non-factor, single position
  - vs Dynamic opponent: rhythmic, random static, random movement
  - vs Enemy agent
- Adding Noise
- Transfer: Mirror, Another rod.

Metrics: speed of ball, time to completion, success rate (opponent type).

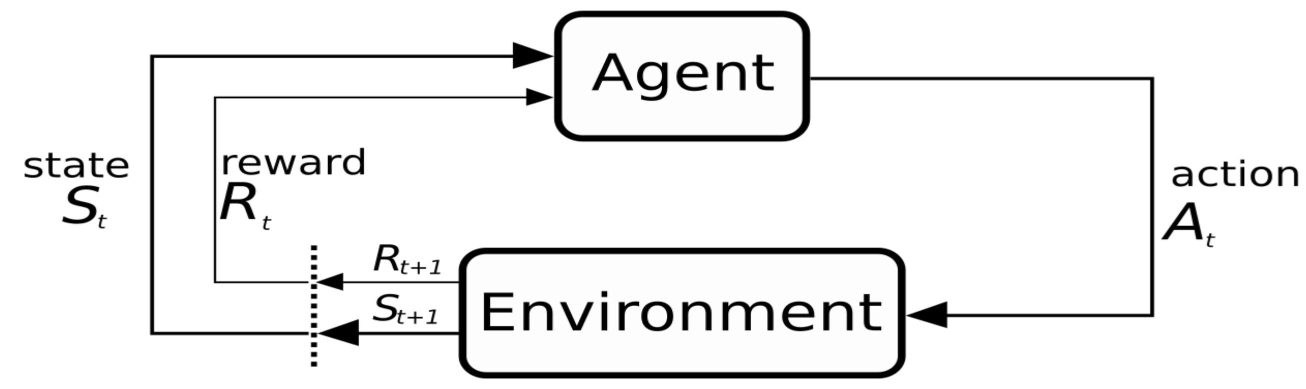
[1] Z. Fan, R. Su, W. Zhang, and Y. Yu. Hybrid actor-critic reinforcement learning in parameterized action space, 2019.
[2] A. Morris and F. Cushman. Model-free rl or action sequences? Frontiers in Psychology, 10, 2019.
[3] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. Proximal policy optimization algorithms, 2017.

Methodology:

MDP: {S, A, T, R, $p(s_0)$, γ}

Model free[2], On-policy

Policy Gradient method: PPO[3]



Reinforcement Learning Setup:

Observations (Normalized)

Only Necessary Rods: (x, θ)

Ball: (x, y)

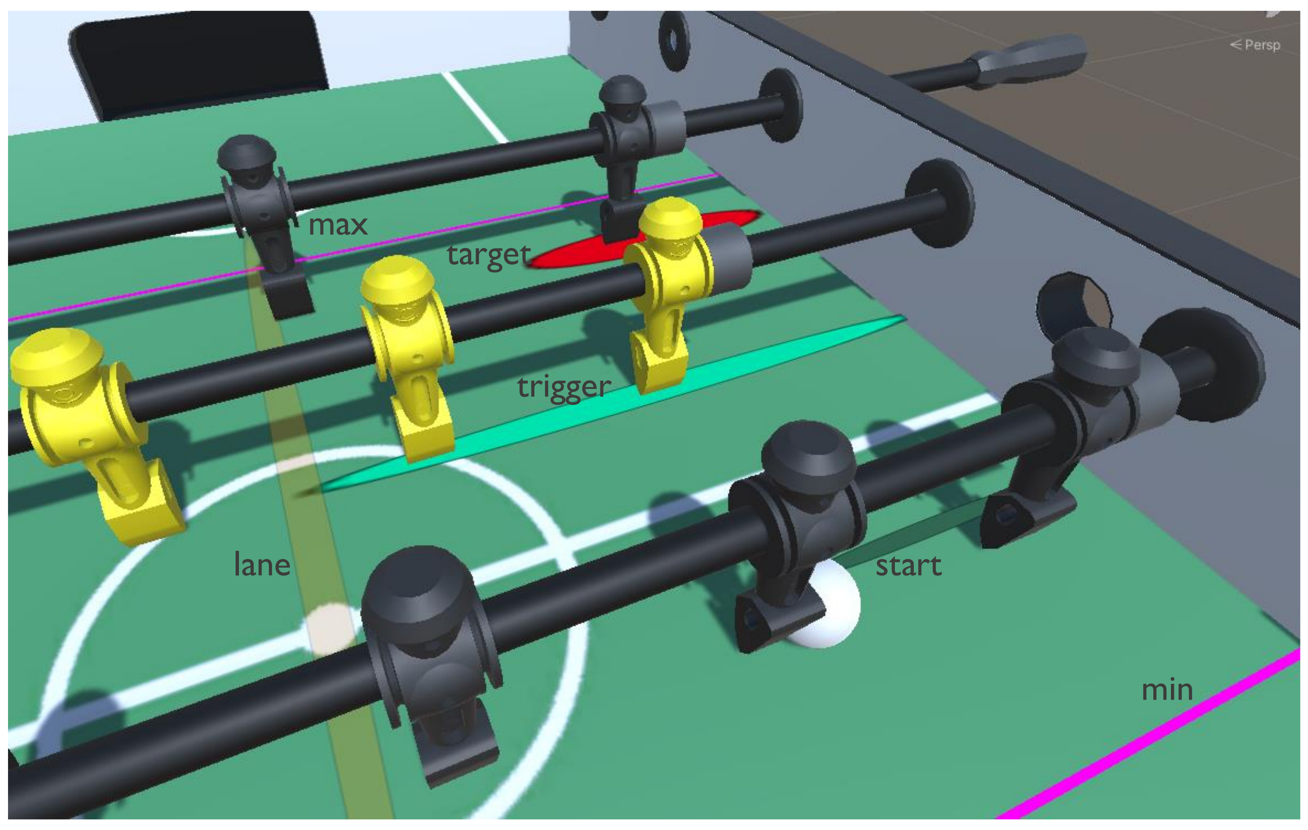
Actions

Translation: Action × Δt × Speed.

Rotation: Action × Δt × Speed.

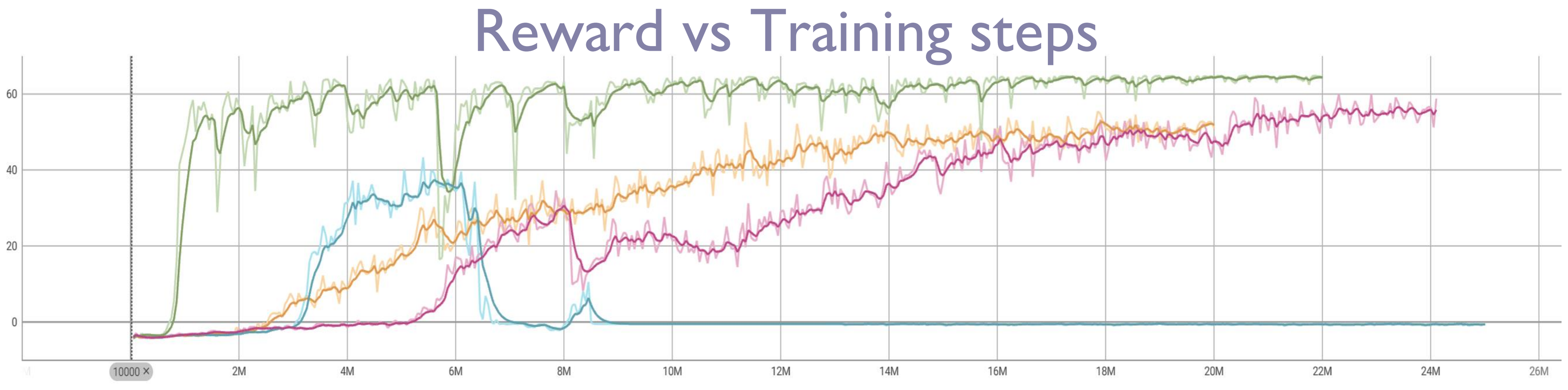
Rewards:

Ball position or collisions.



Results:

Reward vs Training steps



Tick-Tacking:

- Discrete all: Easy to train, Jumpy moves not human-like.
- DCCS: Slower at start, Learns robust moves.
- DCCM: Discrete choice is not helpful for tick-tacking.
- Continuous all: Slower to train, Human-like moves.