ATTRACTION-BASED REINFORCEMENT LEARNING:
A REAL-TIME APPROACH USING TECHNIQUES BASED ON ANIMAL BEHAVIOR

Marcos A. M. Laia, Edimilson B. Santos, Wesley S. Guimarães
Departament of Computer Science
Federal University of São João del Rei (UFSJ)
São João del Rei, Minas Gerais - Brazil

Márcio Mendonça, Rodrigo H. C. Palácios, Janaína F. S. Gonçalves
PPG of Mechanical Engineering - CP/PG
Technological Federal University of Paraná (UTFPR)
Cornélio Procópio, Paraná - Brazil

MOTIVATION AND INSPIRATION

- Machine Learning Challenges:
  - Replicating ideal behavior for tasks.
  - Complex behaviors with fast response.
  - Real-time learning.

- Animals exhibit instinctive behaviors based on their needs:
  - Attraction to food.
  - Repulsion from predators.
  - Pursuit of prey.
  - Evasion.

MOTIVATION AND INSPIRATION

- Simulating animal behavior in virtual environments provides a unique opportunity to design agent-based systems that mimic real-world dynamics.

- These physical rules are useful not only for controlling agents but also for optimizing strategies through machine learning, improving decision-making in dynamic environments.

Games

Real World

OBJECTIVES

- Introduce a machine learning algorithm based on attraction/repulsion using the Unscented Kalman Filter (UKF) to predict and learn opponent behavior in real-time.
- Show the algorithm's effectiveness in dynamic environments, outperforming traditional Q-learning methods, particularly in low-frame conditions.
- Provide a framework for enhancing predictions and facilitating learning in complex, real-world scenarios.

PROBLEM STATEMENT

- Low-Frame Environment Challenges:
  - The low-frame environments, these methods often suffer from reduced accuracy and performance, as they struggle to predict opponent behavior efficiently.

- Need for Improved Approach:
  - There is a clear need for an approach that can combine learning with predictive modeling to optimize both movement and decision-making in real time.

- Parameter Tweaking Issues:
  - The existing techniques for simulating agent behaviors often involve significant parameter tweaking, limiting their practical application in real-world scenarios.

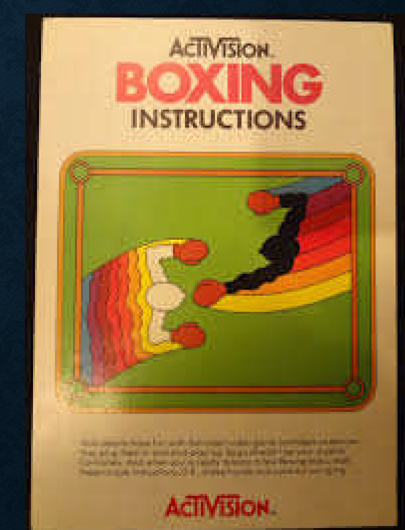
PROPOSED SOLUTION

- Dynamic Learning:
  - This approach allows agents to learn from their environment, adjusting attraction and repulsion parameters to optimize their strategies dynamically.

- Enhanced Predictions:
  - The combination of attraction rules with machine learning enables more robust predictions and learning processes compared to traditional methods like Q-learning.

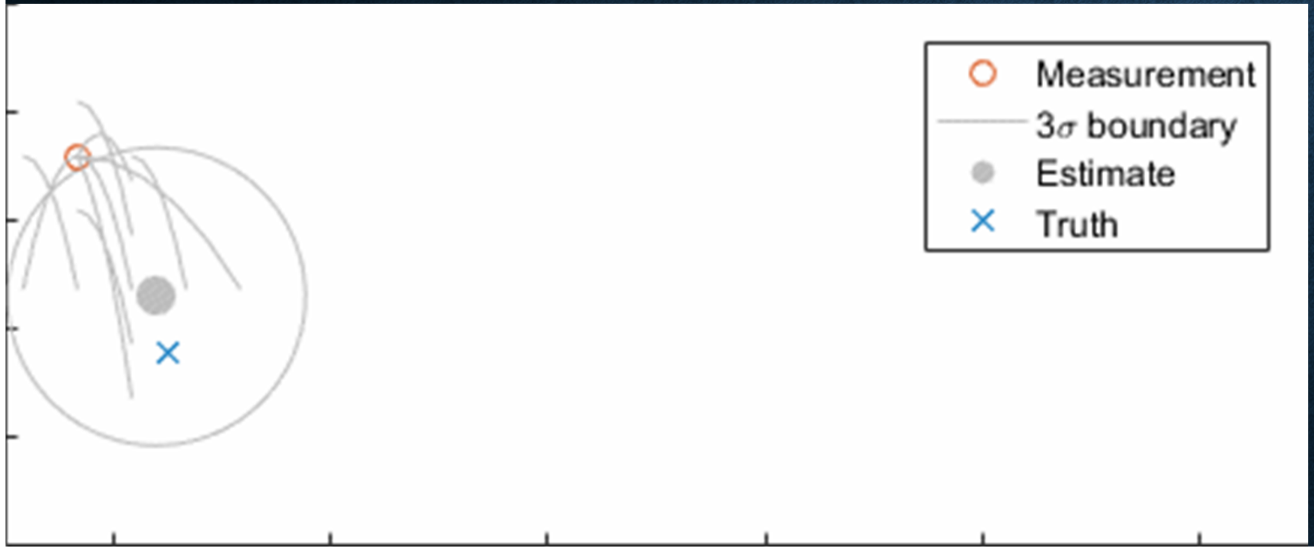
MATERIALS AND METHODS

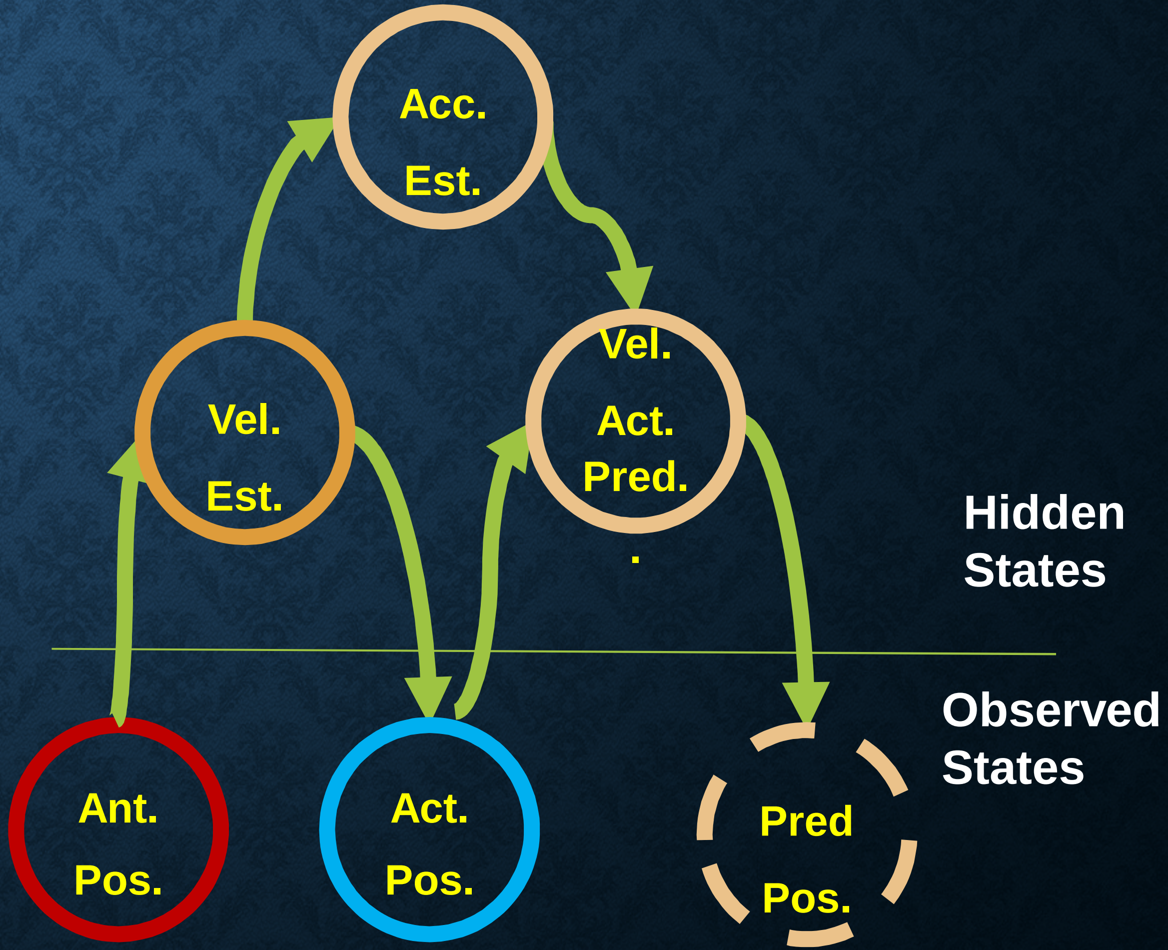
- Atari Emulator provided by OPENAI:
  - Allows the use of games by accessing inputs (buttons) and analyzing the screen output.
  - Enables the simulation of specific scenarios.
  - Provides memory access for reading and writing data for modifications.



MATERIALS AND METHODS

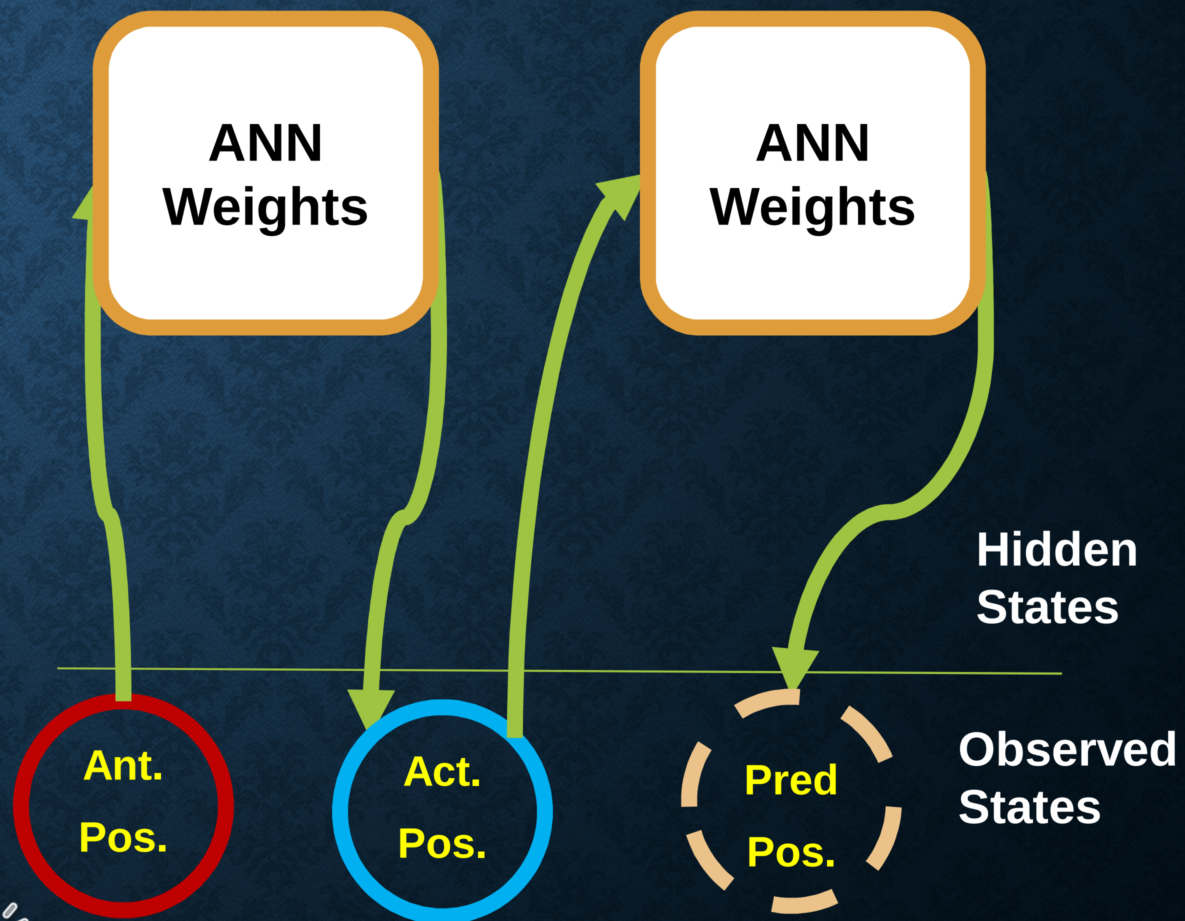
- Discrete Kalman Filter (DKF):
  - Can estimate hidden states, such as velocity and acceleration.
  - Helps eliminate data noise, such as false position measurements.





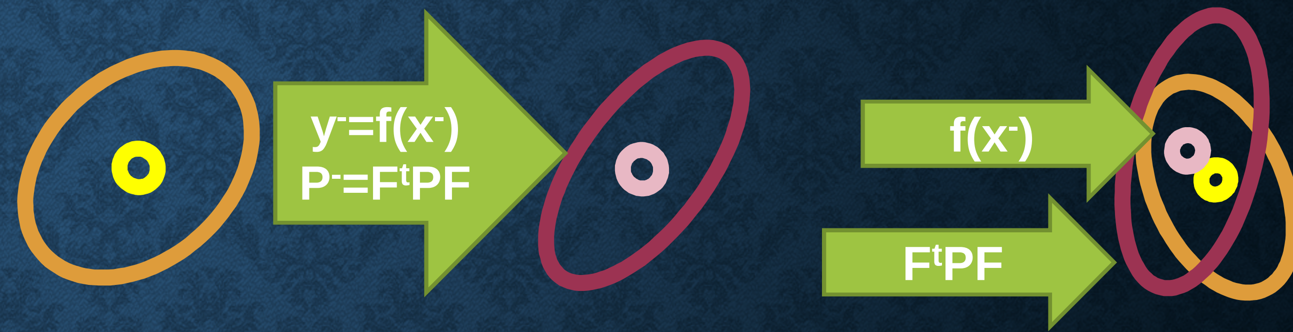
MATERIALS AND METHODS

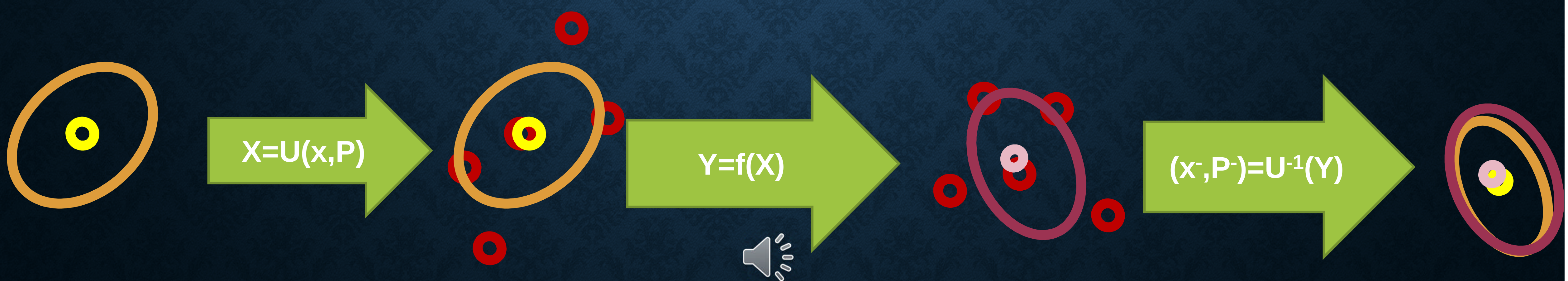
- Extended Kalman Filter (EKF):
  - An extension that uses partial derivatives to handle nonlinearities.
  - When combined with artificial neural networks (ANN), it allows for better state estimation.
  - Can be used to enable an ANN to learn in real-time.



MATERIALS AND METHODS

- Unscented Kalman Filter (UKF):
  - A filter used for nonlinear systems.
  - It includes an improvement with sigma points for better mapping of the intrinsic variances of the states.

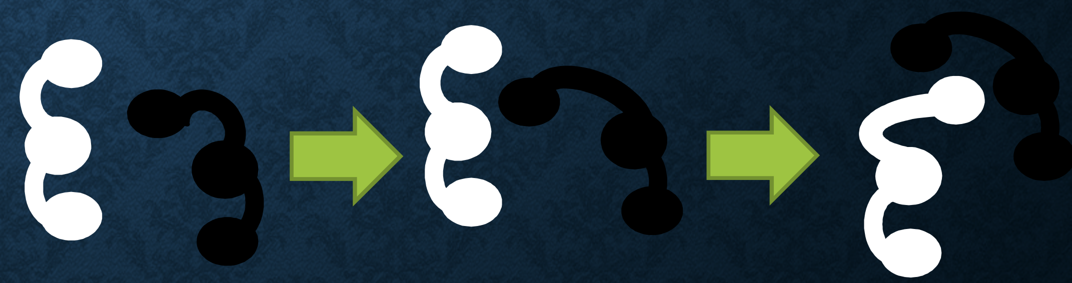




ATTRACTION-BASED MODEL

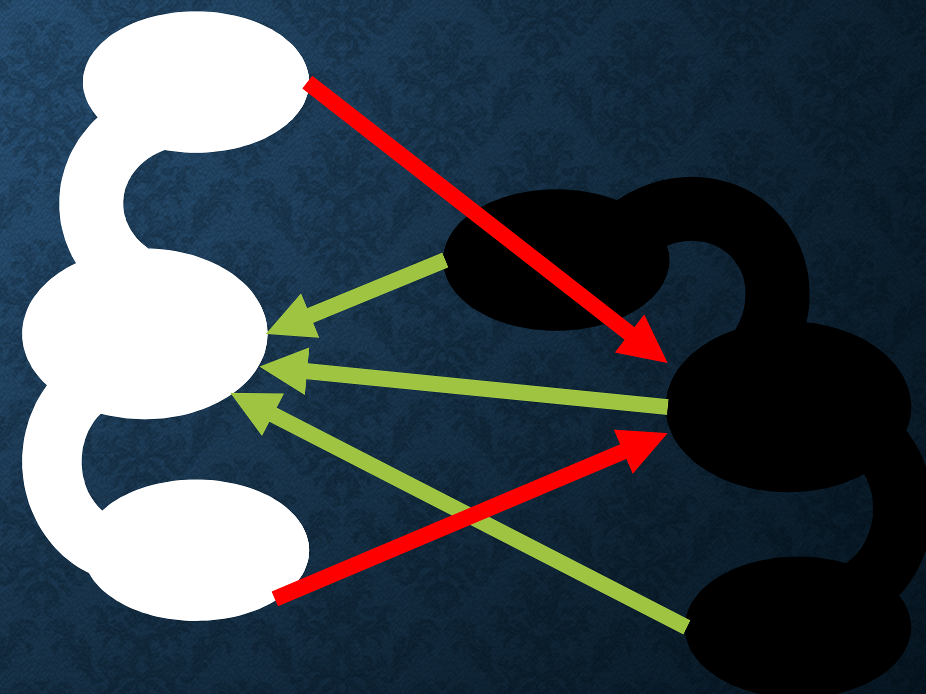
- Model based on attraction and repulsion techniques:
- The player's fighter is attracted to the CPU-controlled fighter.
- The fighter experiences repulsion from the opponent's glove.
- At a certain distance from the player's gloves and the CPU's head, the player throws a punch.

```
Player_pos=Attraction(Player_pos,CPU_pos,K_1);
Player_pos=Attraction(Player_pos,CPU_G_R_pos,-K_2);
Player_pos=Attraction(Player_pos,CPU_G_L_pos,-K_2);
```



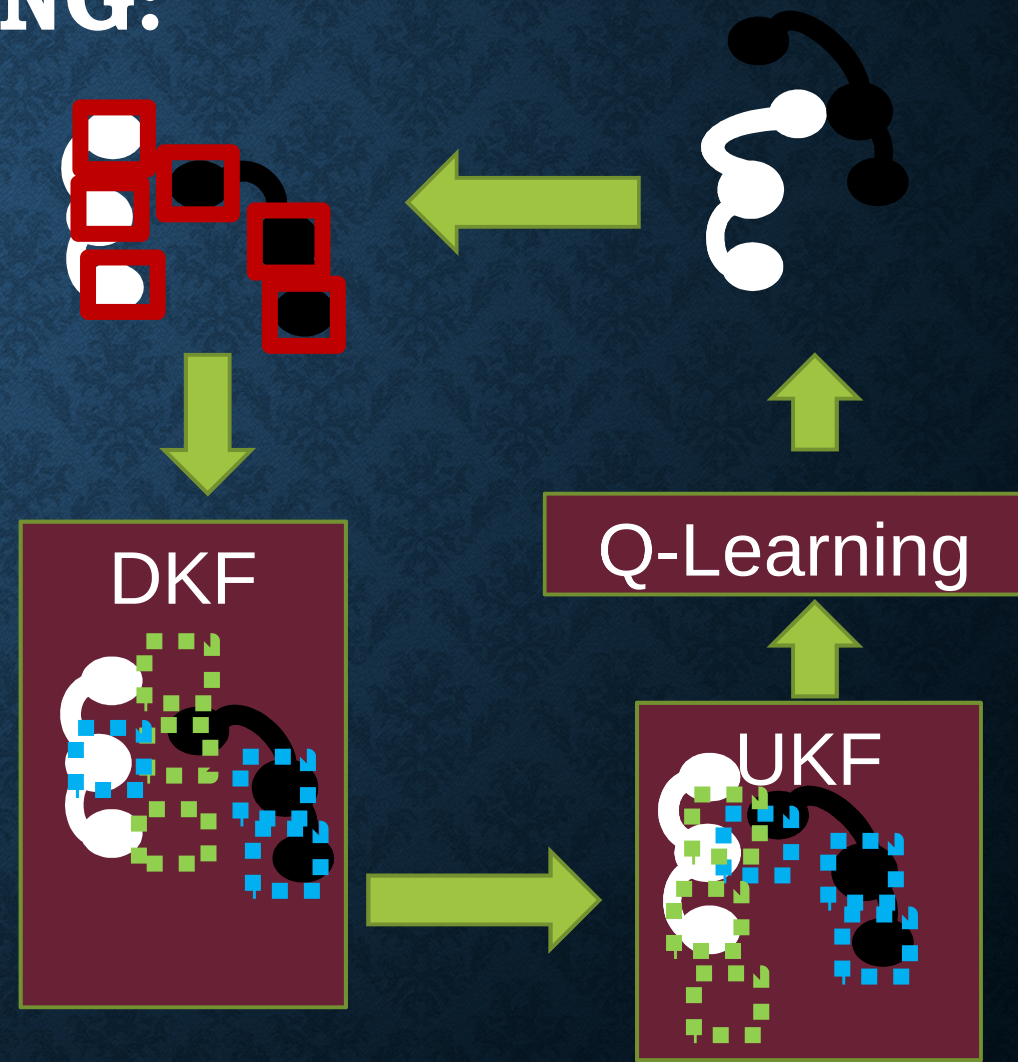
ATTRACTION-BASED LEARNING

- Learns the degrees of repulsion and attraction between certain sprite positions using the DKF and UKF.
- It is done in real-time.
- The reaction time starts from the moment the coefficient values are assigned.



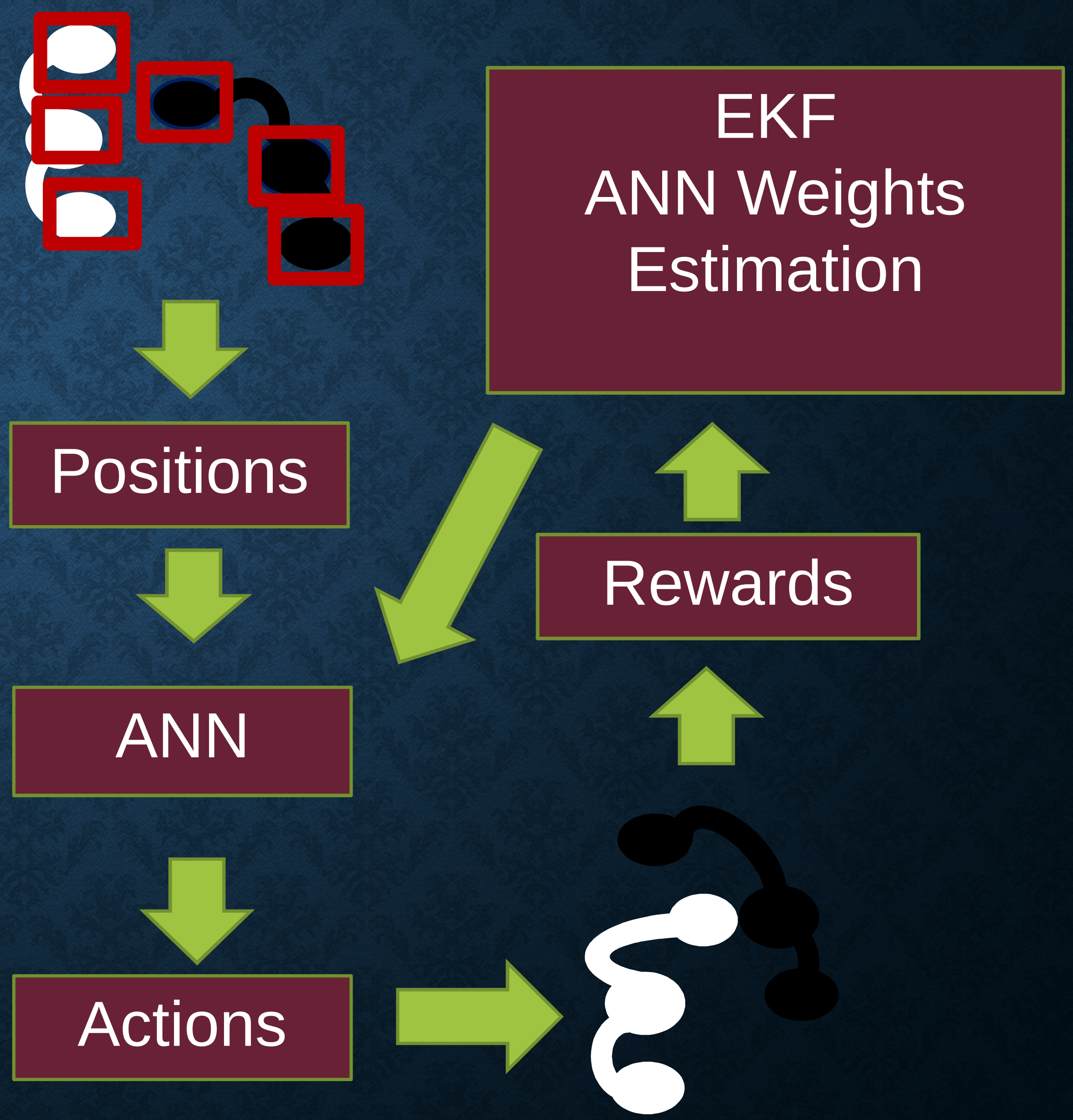
ATTRACTION-BASED REINFORCEMENT LEARNING:

- We perform a linear prediction of the future position of the sprites using the DKF.
- We use attraction-based learning to better determine the sprites' positions in future steps.
- The learning process is enhanced by adding an objective through rewards.

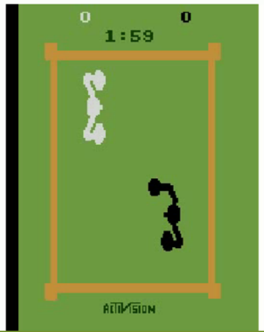


NOT TOO DEEP Q-LEARNING USING EKF

- We use the EKF to train the neural network in real-time.
- It replaces the heavy processing used in Deep Q-Learning by already having the sprites' positions.
- The neural network has fewer neurons and fewer layers, making it less complex.
- This allows training with the lowest possible error.



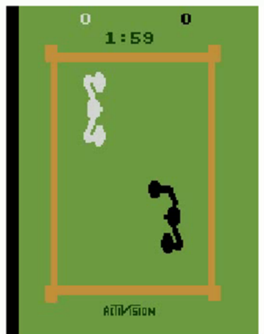
RESULTS



UK
F

Q-Learning

EKF



UKF+Q-Learning

UKF + EKF

Algorithms	Score	Remaining Time	Total Rewards
Programmed	KO x 74	34 s	45,422
UKF	79 x 78	0 s	27,670
Q-learning	KO x 97	3 s	33,361
EKF + RN	24 x 58	0 s	-18,241
UKF + Q-learning	95 x 96	0 s	30,871
UKF + EKF	13 x 59	0 s	-42,262

CONCLUSIONS

- Integrated Approach: The integration of attraction and repulsion rules with the Unscented Kalman Filter presents a promising solution for real-time decision-making in dynamic environments.
- Enhanced Learning: The proposed method enhances the learning process by refining predictions about agent movements and behaviors, leading to more robust strategies.
- Improved Reliability: Compared to traditional Q-learning and EKF-based approaches, the UKF offers a more reliable and adaptive framework, particularly in low-frame-rate conditions where fast, real-time decisions are required.
- Demonstrated Potential: The results demonstrate the potential of combining physical interaction models with machine learning to improve both prediction accuracy and agent performance in complex environments.

FUTURE WORKS

- Model Refinement:
  - Future improvements will focus on refining the attraction and repulsion models to incorporate more complex behaviors and interactions, such as environmental factors and multi-agent dynamics.
- Real-World Testing:
  - The next steps include testing the algorithm on video data of real-world players to enhance the model's ability to predict human behavior.
- Sports Analytics Application:
  - Applying this algorithm to other fields, such as sports analytics, could provide new insights by analyzing player movements and strategies in real-time
- Robotics Exploration:
  - Further exploration into robotics is planned, particularly in areas such as real-time

AUTHORS THANKS

- The authors would like to thank the Federal University of Technology – CP/PR and Federal University of São João del Rei for the financial support provided for this research.

Universidade Federal de São João del-Rei

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

marcoslaia@ufsj.edu.br