

Systems Analysis & Design

**REINFORCEMENT LEARNING
AGENT FOR GOOGLE
RESEARCH FOOTBALL WITH
MANCHESTER CITY F.C.
COMPETITION**

Universidad Distrital Francisco José de Caldas

Juan David Escallón Guzmán, Juan Diego Lozano Luna, Jorge Eduardo Muñoz Gómez

INDEX

- Competition Context
- Systematic Approach
- Recognizing Requirements
- Designing a Modular Solution
- Some Experimenting and Emergent Behaviors.
- Outcomes And Conclutions

COMPETITION CONTEXT

**ONE PLAYER AT
TIME**

**NO REWARD
FOR HIGH
SCORE
DIFERENCE**

**LEFT TEAM
PERSPECTIVE**



**MATCH
CONSIDERATIONS**

**NO SWITCH
SIDES**

3000 STEPS

**NO FINAL
PENALTIES +
NO EXTRA TIME**

OPEN-SOURCED SIMULATION ENVIRONMENT

- Preseted scenarios to train
- Easy simulation of the competition
- Highly modifiable
- Easy API implementation



AVAILABLE RESOURCES

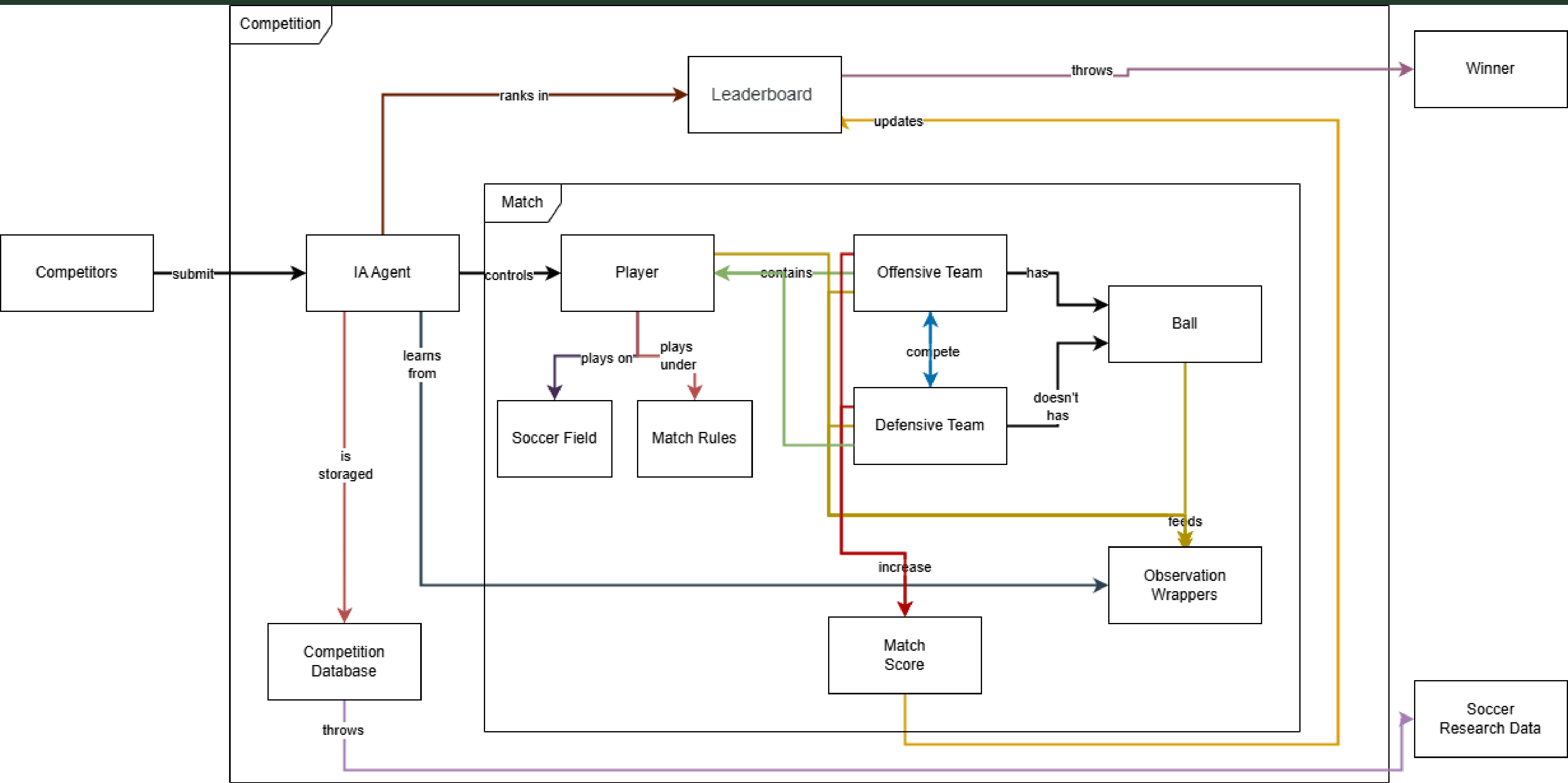
HUGE VARIETY OF OBSERVATIONS FORMAT

Includes game information like:

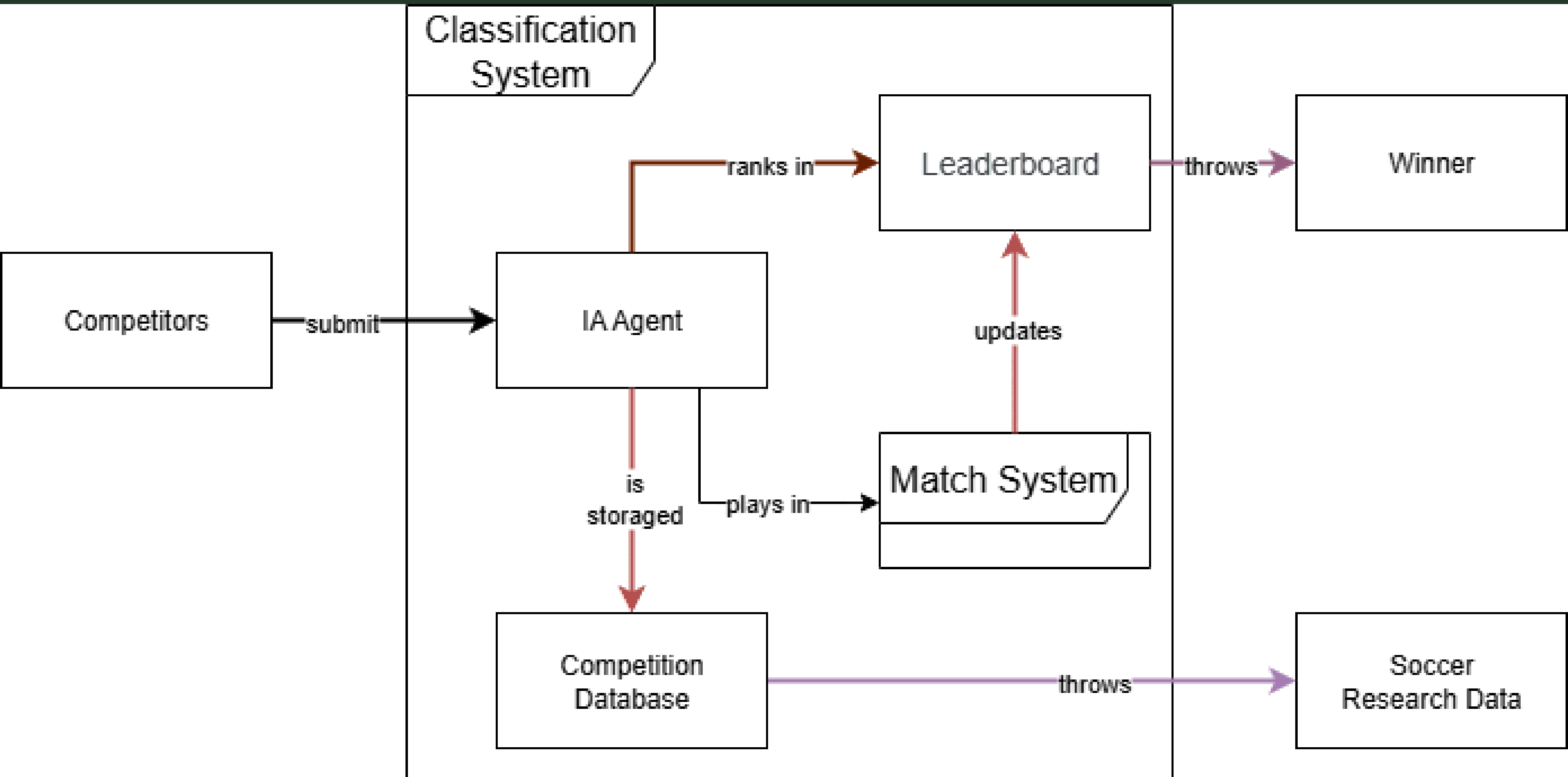
- Position, velocity, and fatigue of all players
- Ball position and possession
- Current match score
- Current game mode (corner kick, throw-in, etc.).

SYSTEMATIC APPROACH

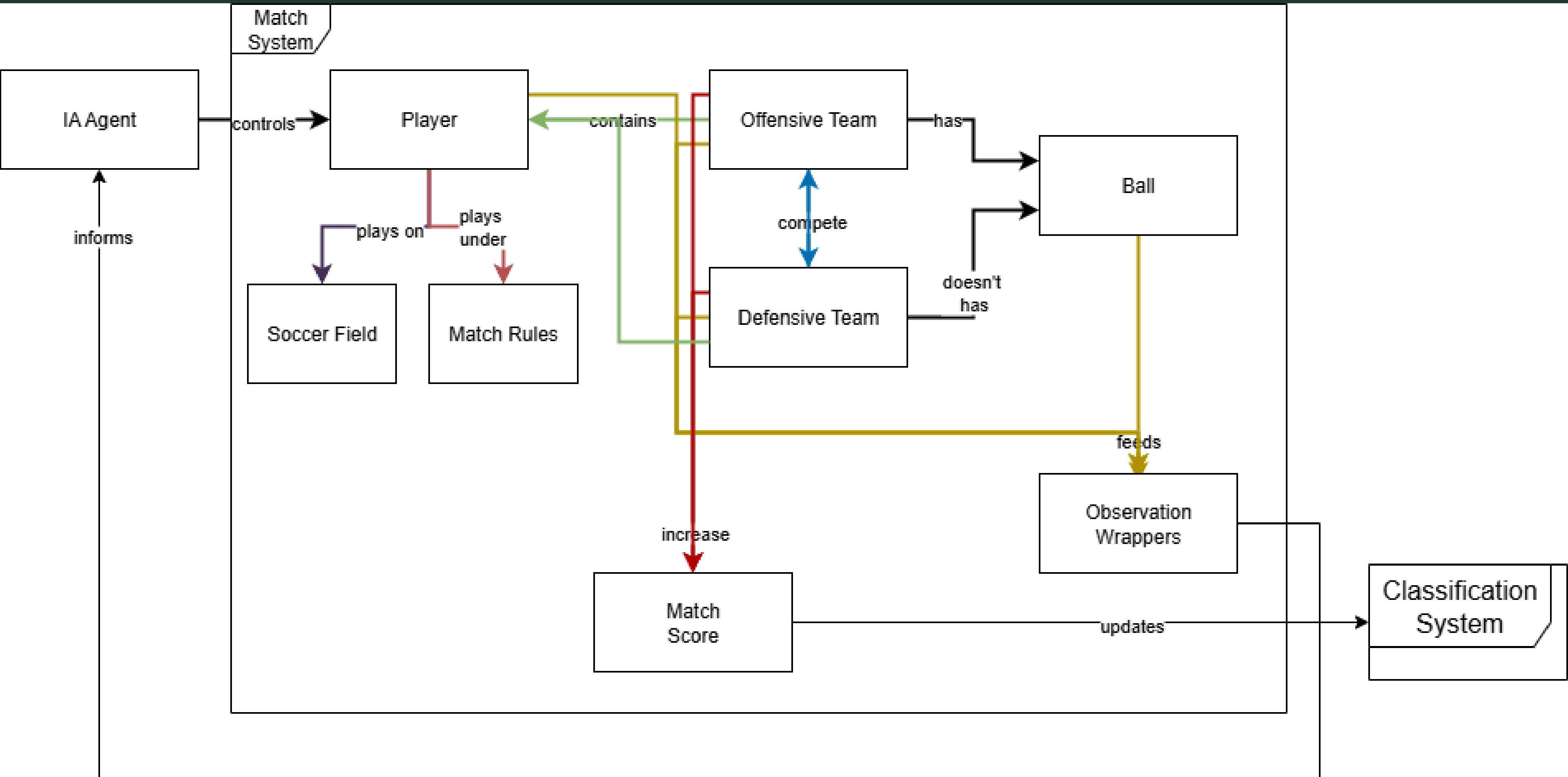
FULL COMPETITION SYSTEM



CLASSIFICATION SYSTEM



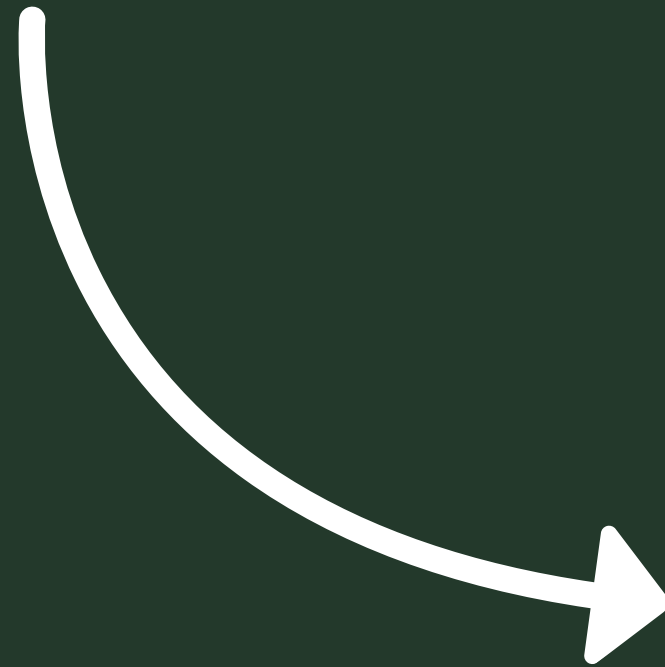
MATCH SYSTEM



CHAOTIC BEHAVIORS ON COMPETITION

RECOGNIZING REQUIREMENTS

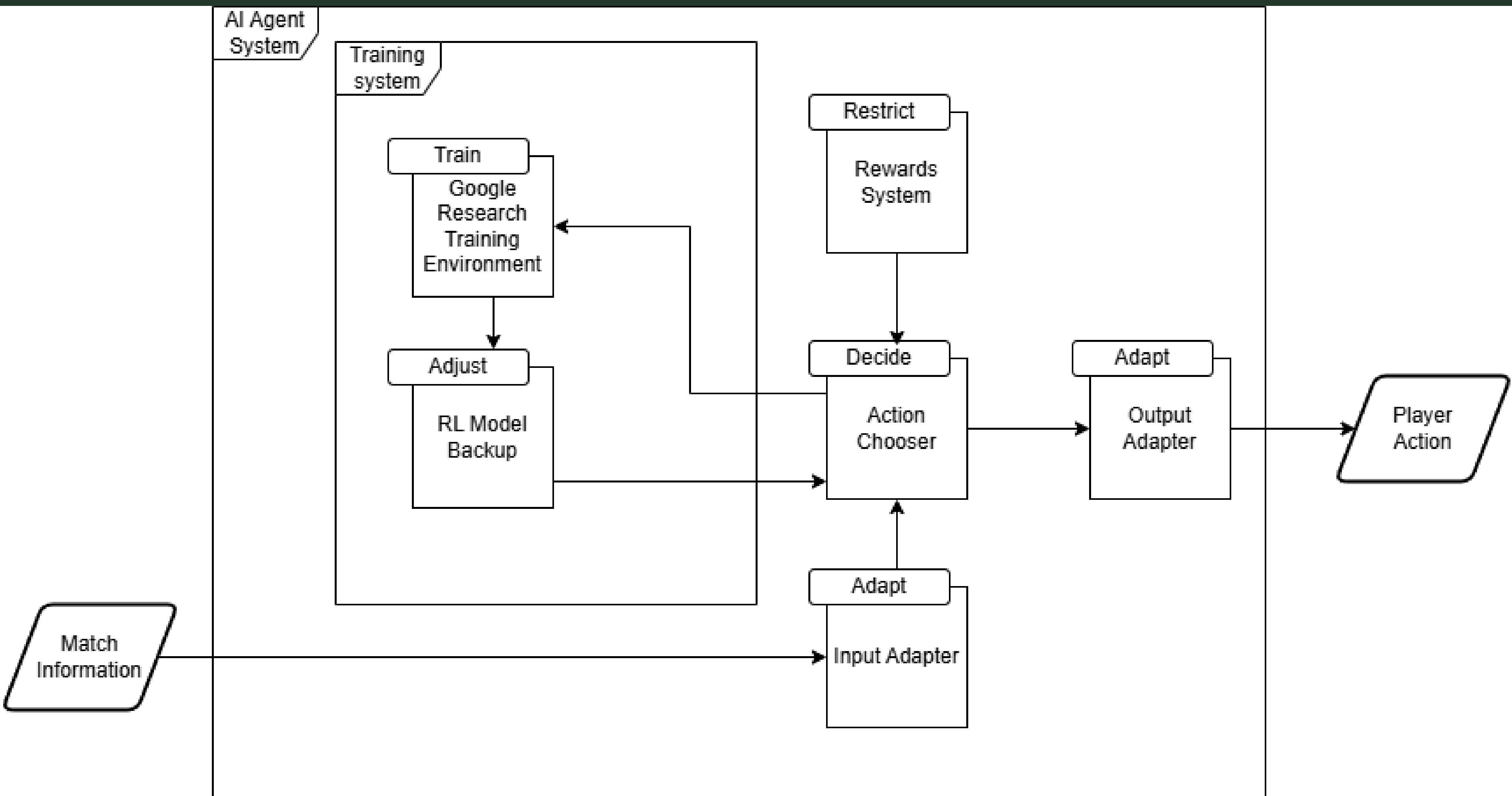
**GAME
INFORMATION**



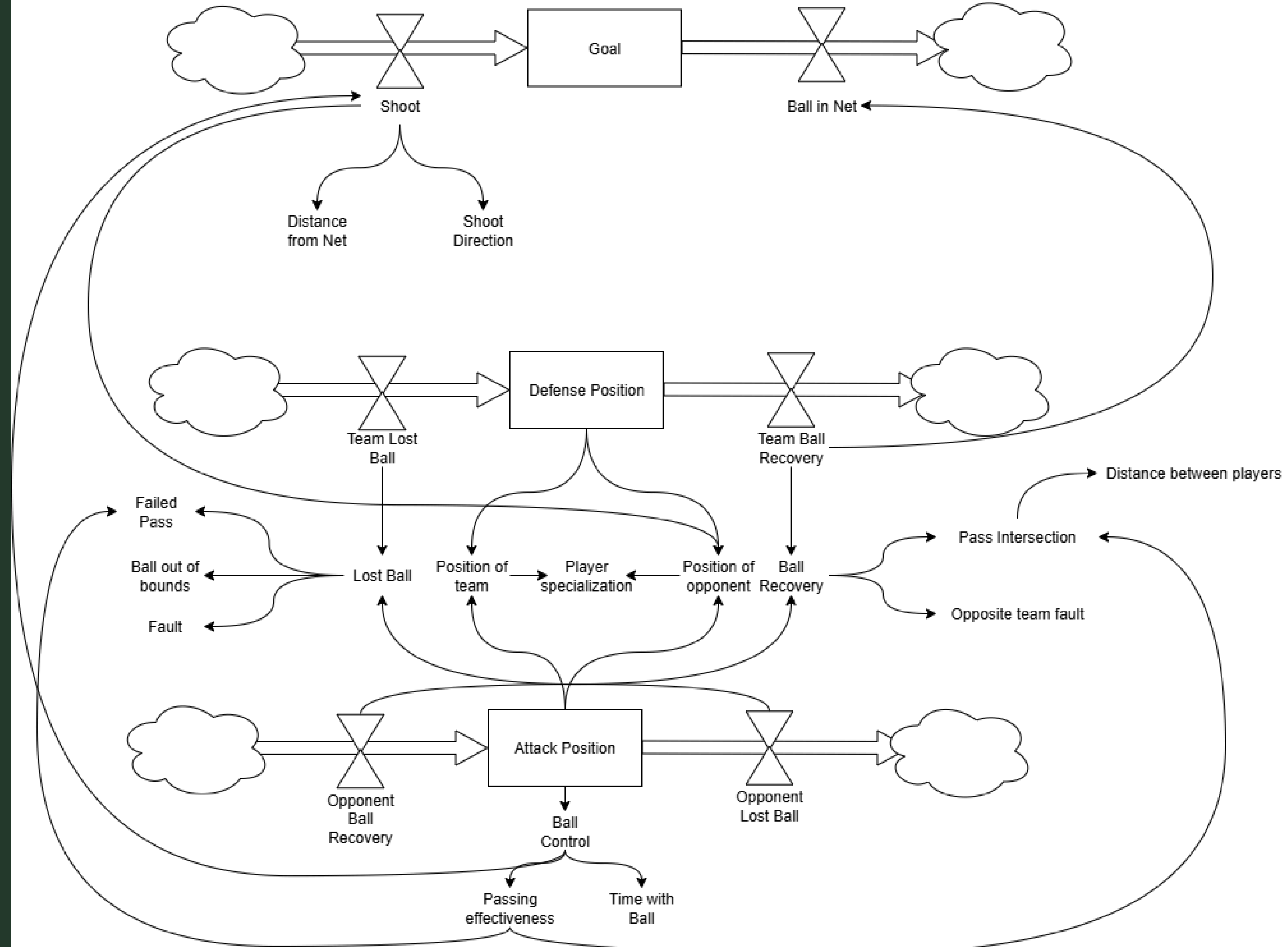
**1 FROM 19
ACTION TO THE
ACTIVE PLAYER**

- Correctly interpreting the data provided by the simulation environment, regardless of the format in which it is represented.
- Make precise and adaptive decisions, always complying with the rules of the game.
- Confront possible chaotic learning
- Consider the answer format solicited by the competition

DESIGNING A MODULAR SOLUTION



**¿WHAT SHOULD BE
CONSIDER IN REWARDS
SYSTEM?**



SOME EXPERIMENTING AND EMERGENT BEHAVIORS

- Emergent behaviors on training process.
- Different outcomes depending on training scenarios
- Several modifications to the reward system weights

OUTCOMES AND CONCLUTIONS

- **Flexibility in designing** the reward system, as the model translates program information into a format understandable by a general audience, making the reward structure easier to analyze and refine.
- Not a meaningful difference in behavior based on the **player's position** or specialty, which was one of our main strategies.
- **Splitting the model in two**, with offensive and defensive action chooser could bring better results.

THANKYOU