Game Plan:

Each player has:

- a location attribute in the form of a coordinate (x,y)

- whether it has the ball in the form of a boolean

- every second, each player decides how s/he should act

Player actions include

- intercept

- run

- expressed as a vector, with the magnitude being the speed and the direction

being the direction

- pass

- shoot

The ball has:

- a location attribute in the form of a coordinate (x,y)

- the speed and direction it's going in the form of a vector

Training Plan:

- going near the ball = reward

- intercepting the ball = bigger reward

- having the ball taken from you = punishment

Have a transition function that takes an action, a state (everything happening at the moment; a snapshot in time), a specific player this is intended for, and returns a new action for this specific player

We plan on doing policy iteration to train the agents. Not sure how though.

We should have a while loop that keeps calling the transition function, which requires a reward function.

Each player should have his own policy. Then there’s a separate reward function.

**Important: 1v1 should be our first milestone, maybe even only one player only first!!**

Right now, we’re considering only have two players – one player on each team. When there’s only one player on each team, the only actions are run and shoot. And the probability of reaching a new state is always 100% for now.

Low reward for shooting from afar, even if it gets in. High reward for getting it in a short distance.

R(s, a, s'){

match a with

| ball\_run -> return whatever

| ball\_shoot -> return whatever2

| no\_ball\_run -> return whatever

...

}