# Initial Strategy:

CNN Design:

One matrix of size board size for:

1. Units positions
   1. Friendly units
2. Relic node positions
   1. Relic node points
3. Asteroid Positions?
4. Nebula Positions?
5. Board energy levels?

Somehow need to encode the unit energies, how to do this?

Final embedding size 512?

Actor-Critic Desing:

Different heads:

Actos is dense layer with 512? In en actions out

1. Movement actions 5
2. Sap action how is this encoded?
   1. Need to specify wether to sap and then where

Critic is dense layer with 512? With 1 output (the value of the current state)

Reward Function:

R = ?

Reward exploration: unseen tile see + 1

Reward getting points: points + 10

Sapping: + 5

RL method – Training

PPO – training

Multi-agent – Training

Have the same agent for both sides initially to learn to play the game.

Later Restart training from a different stochastic process and force to learn different things via reward function

E.G reward sapping more heavily