Hyperparameters and Neural Network Configuration for all experiments (Across all 3 algorithms)

1. **Constrained Cooperative Navigation(Number of Agents = 5):**
   α= 3
   β(t) = 0.00001
   Seed = 0
   Episode length = 25
   Steps before an update = 50
   Number of updates(performed at each update)= 2
   Number of rollouts= 12
   Buffer length = 1e6
   Batch Size = 1024
   Tau = 0.001
   Gamma = 0.99
   Learning rate for all policies and critics = 0.001
   Hidden Dimension for all policies and critics = 128
   Activation Function = Leaky Relu
   Total number of iterations = 200k
   Temperature coefficient = 0.01

2. **Constrained Cooperative Treasure Collection(Number of Agents = 8)**
   α1= 12
   α2 = 0.2
   β(t) = 0.00001
   Seed = 0
   Episode length = 100
   Steps before an update = 100
   Number of updates(performed at each update)= 4
   Number of rollouts= 12
   Buffer length = 1e6
   Batch Size = 1024
   Tau = 0.001
   Gamma = 0.99
   Learning rate for all policies and critics = 0.001
   Hidden Dimension for all policies and critics = 128
   Activation Function = Leaky Relu
   Total number of iterations = 100k
   Temperature coefficient = 0.01