Learn latent structure model to assign roles to agents.

For each role, train its own policy model (supervised learning).

Alternatingly repeat the two processes until convergence.