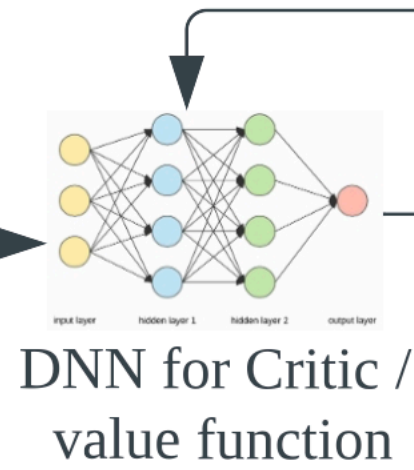
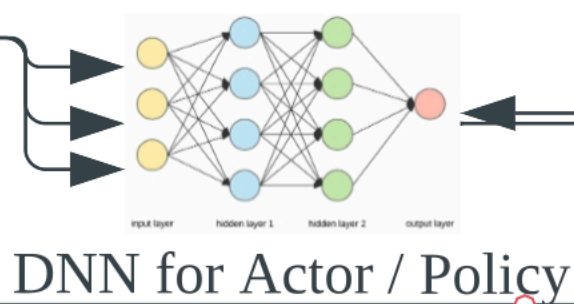


Agent Tank



compute advantage estimates A , and optimize objective L .

State, S



Observer State, S

Actions, a

Environment