

ロールアウト
ポリシー

教師あり学習による
ポリシーネットワーク

強化学習による
ポリシーネットワーク

価値
ネットワーク

ニューラルネットワーク

p_π

p_σ

ポリシー勾配

p_ρ

v_θ

データ



人間のプロ囲碁棋
士の対戦棋譜

セルフプレイ
自動生成による棋譜