

強化学習を用いた 量子回路の最適化

Team: QuRL

Team Member: @king-p3nguin, @Ojun52, @Naoki0824

Team Mentor: @SoshunNaito, @qphkk

量子回路の最適化

意義

- ・誤り耐性の向上、ノイズの影響軽減
- ・計算コストの削減と高速化

設定

簡単のためCNOTゲートのみからなる回路を考え、CNOTゲートを適用することで、回路内でCNOTゲート同士が打ち消し合い、最終的に恒等回路にすることを目標とする。

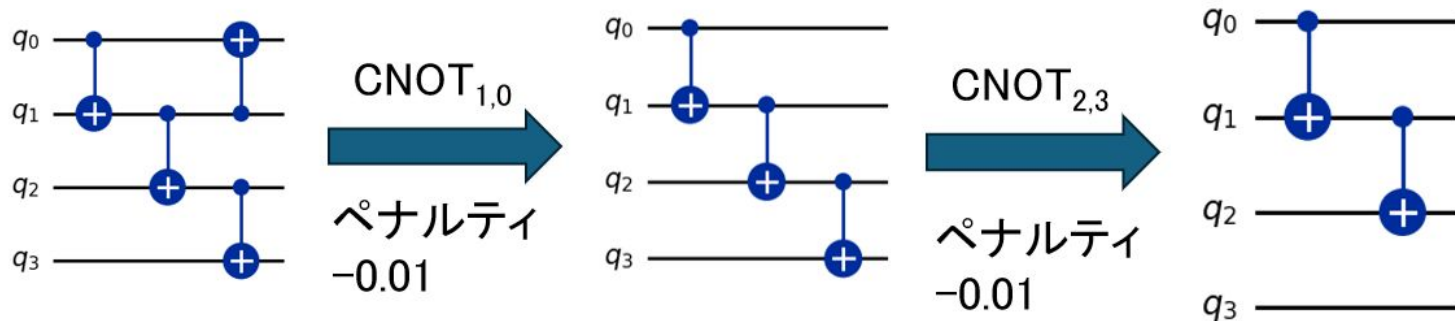
→その逆操作で量子回路は最適化される。

→強化学習を用いる

強化学習による最適化の試み

強化学習: 「報酬」という数値を最大化するような行動を、試行錯誤で学習する仕組み。

先行研究における報酬設計: CNOTを置くたびに微小なペナルティ。ゴールに到達したら大きい正の報酬。

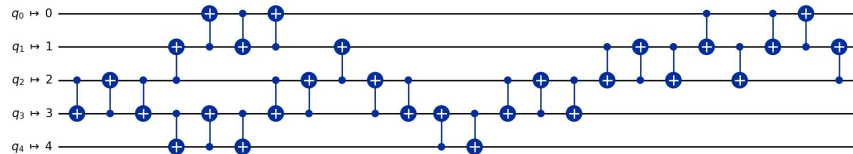
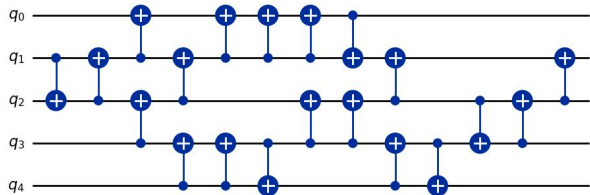


強化学習を用いてCNOT数を最適化

Qiskitの関数*

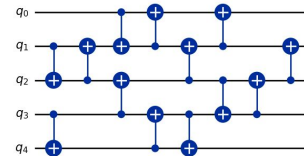
*Qiskitのsynth_cnot_count_full_pmh関数
+ coupling mapを固定してtranspile

ランダム生成したCNOTの
みの量子回路



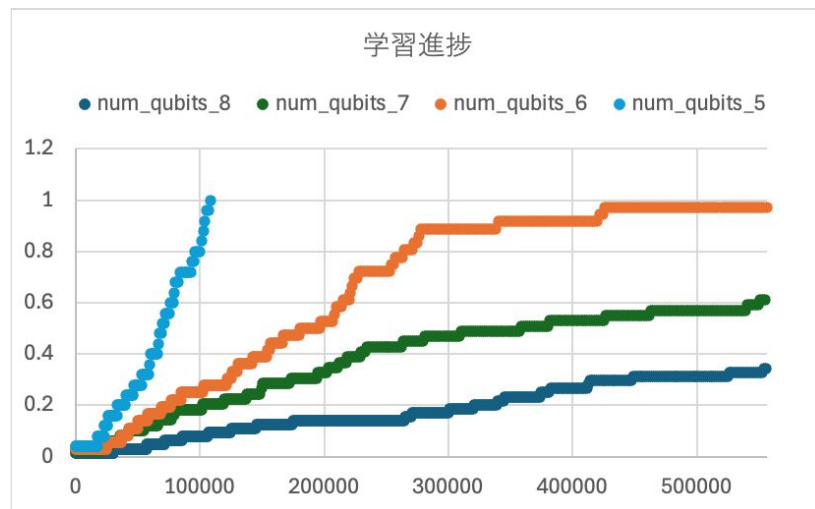
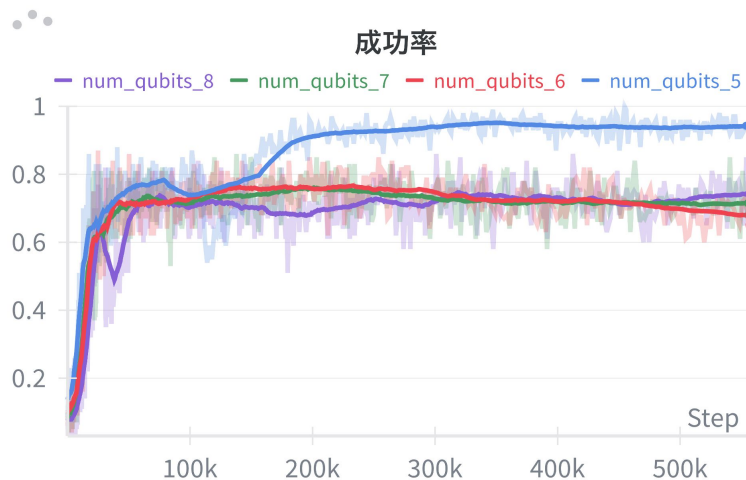
CNOT数: 28, depth:24

強化学習モデル



CNOT数: 13, depth:8

量子ビット数の増加の影響



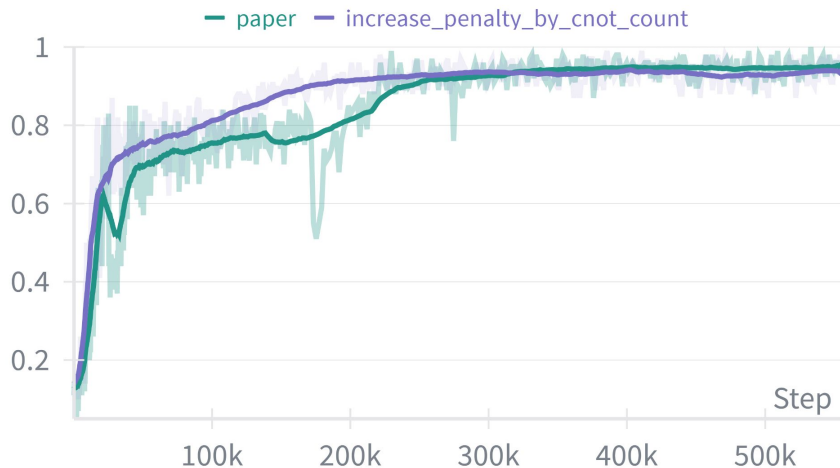
- 量子ビット数を5 → 6 → 7 → 8と増加
- 量子ビット数が増加すると学習速度が遅くなる
→ 学習が速くなるような報酬関数が必要

報酬関数のデザイン

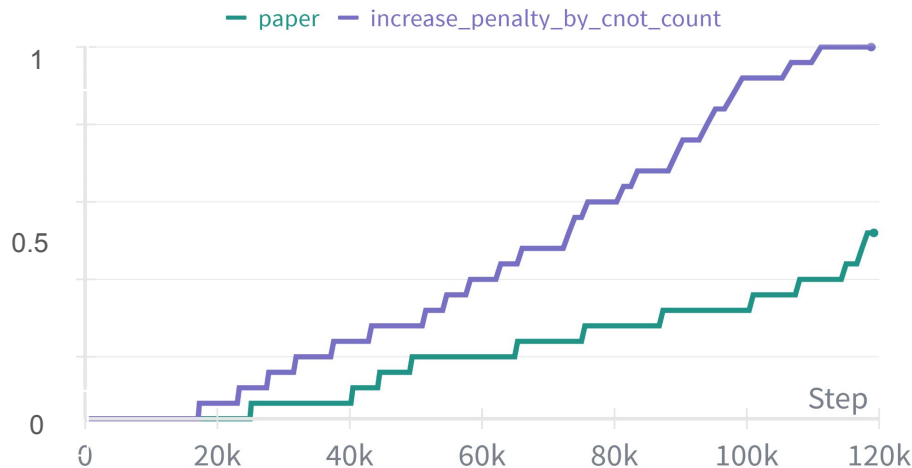
論文: $\text{penalty} = -0.01$ (常に一定)

試してみた手法: $\text{penalty} = -0.01 \times \text{number of CNOTs}$ (ゴールに近づくほどゲート数削減の圧力を強める)

Success rate



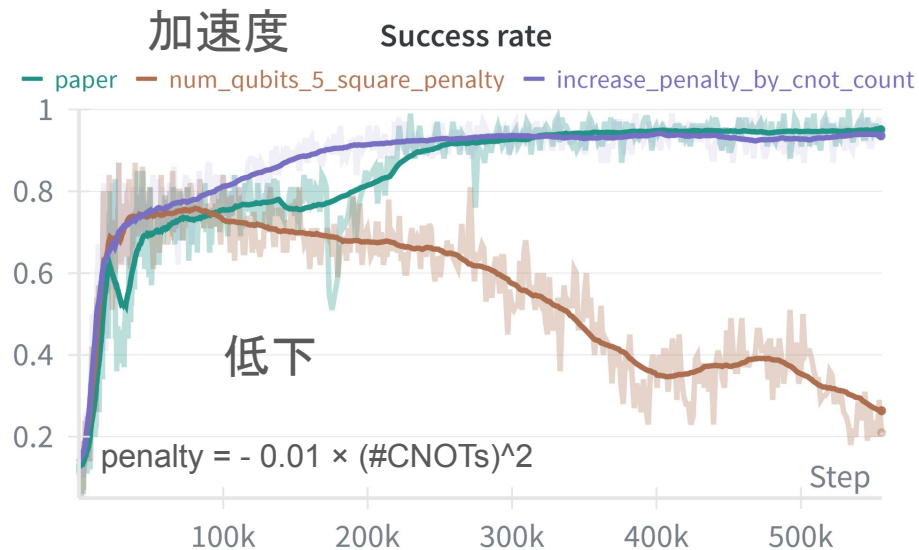
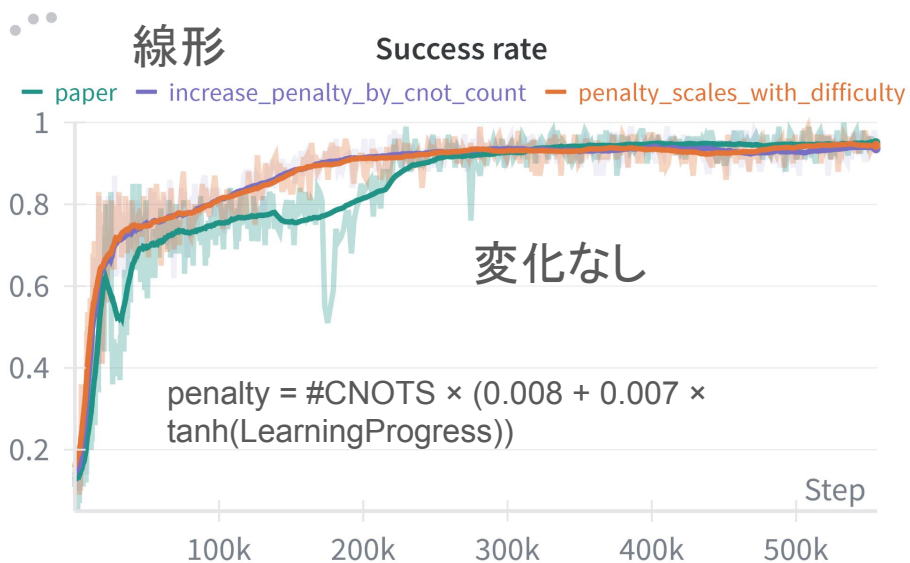
Learning progress



→より上手にゲート数削減の圧力を変化させられないだろうか？

報酬関数のデザイン

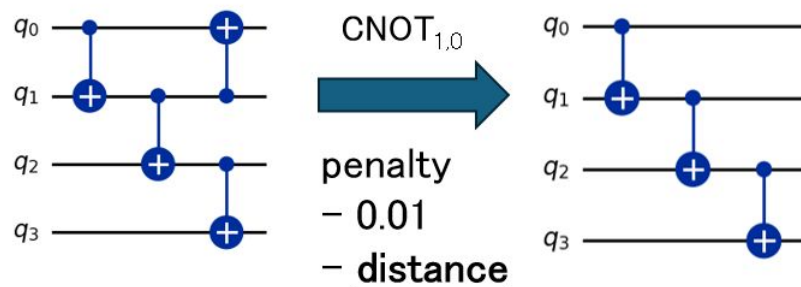
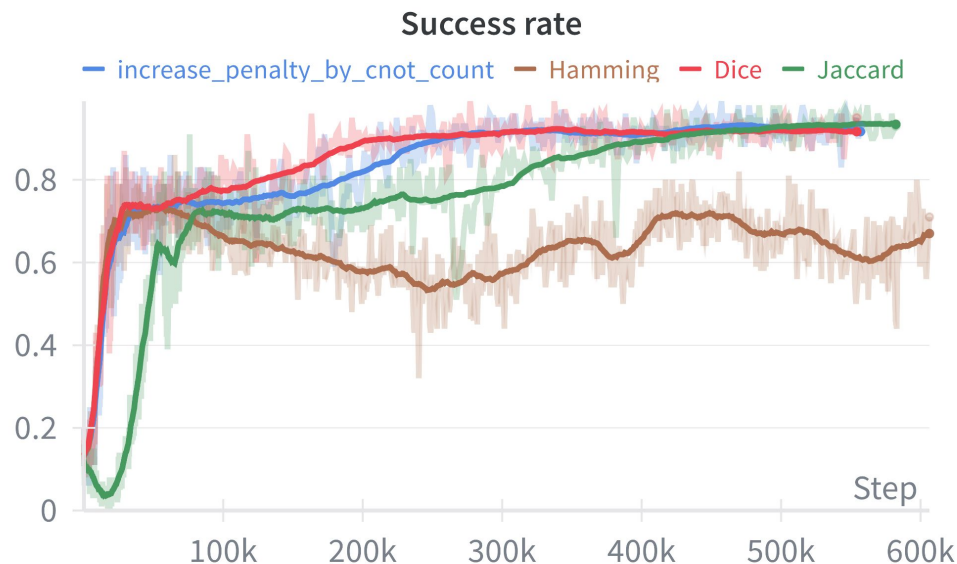
線形/加速度的にpenaltyを増加させた場合を比較。



学習の進みに応じたペナルティのスケールでは大幅な改善はみられない。
→ 他の変数に応じたスケールの探究をすると、「良い」報酬関数の可能性。

距離に依存した報酬関数のデザイン

現在の回路とゴール時の回路を行列表現→距離に応じて報酬を与えるように設定



Hamming距離 Jaccard係数 Dice係数

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad \text{DSC}(A, B) = \frac{2|A \cap B|}{|A| + |B|}$$

大幅な改善は見られなかったが、距離依存性はある→「良い」距離の存在可能性

まとめ

- チームメンバーで様々な独創的アイデアを出し合い、論文とは異なる多彩なアプローチで強化学習モデルを開発した
 - CNOT追加によるペナルティの付与の仕方の変更
 - ゴールとの距離に応じたペナルティの付与
 - 他にも、depthの考慮、制約付き強化学習...など
- 今回試した報酬関数の中には、論文の手法より学習速度が速いものも！
 - コミュニティの中で幅広く利用されることを期待
- 量子回路の最適化は量子計算の高速化に直結するのでインパクトは大きい
- 今回の手法を超えて様々な強化学習モデルや量子回路への拡張を探求したい