

量子回路設計に向けた量子部分観測マルコフ決定過程(Q-POMDP)手法 の開発

一部分観測強化学習を活用した量子回路設計

1. 背景

量子力学の原理に基づき計算を行う量子コンピュータは、従来のコンピュータを超える計算能力を備えたコンピュータとして注目されている。実際、因数分解など、特定の問題を効率的に解決するアルゴリズムが提案されている。しかし、量子コンピュータで特定の課題を解くための量子回路を設計することは容易ではなく、様々な研究がなされてきた。

強化学習は、環境とのインタラクションから自律的に学習を行う学習手法であり、ロボット制御などの幅広い分野で圧倒的に成功している手法である。強化学習が複雑な制御問題に対して解を出せる可能性があるということで、量子回路の制御に強化学習を適用しようとしている研究が行われ始めている。

強化学習では通常環境の情報を完全に獲得できるという仮定(マルコフ決定過程)をおくが、現実の問題ではノイズなどの影響により完全に情報を取得することは難しい。このような場合問題は部分観測マルコフ決定過程(POMDP)によって問題をモデル化する必要がある。

量子系に強化学習を適用する場合、測定を通して得られる情報は部分的な情報のみであることから POMDP として問題を考えなければいけない。実際 POMDP を量子に拡張した手法として量子観測マルコフ決定過程(QOMDP)が提案されている。QOMDP は、POMDP における信念状態を量子状態に一般化したものとして提案されたが、具体的なアプリケーションは提案されていない。

2. 目的

本プロジェクトでは、量子観測マルコフ決定過程(QOMDP)におけるプランニングアルゴリズムを提案し、提案したアルゴリズムを用いて量子回路設計問題を解くことを考える。

本研究では具体的なアルゴリズムとして、POMDP でのプランニングアルゴリズムを参考に、QOMDP におけるプランニングアルゴリズムを提案する。厳密なプランニングは、計算量的に以下の難しい点を持つ。行動・観測の時系列がタイムステップに対して指数的に増えてしまうという問題、ヒルベルト空間に関して和集合を取ることができないという問題、ヒルベルト空間の次元は量子ビットに対して指数的に増えるという問題である。本プロジェクトではこれらのうち 1 つ目と 2 つ目の問題を解決するプランニングアルゴリズムを提案する。

次に、提案したアルゴリズムを用いて量子回路設計問題を解くことを考える。回路設計問題の具体例として状態生成、基底状態探索という二つの問題を挙げる。状態生成に関しては目標とする状態として Bell 状態と GHZ 状態を設定する。また、基底状態探索に関しては、水素分子、水素化ヘリウムイオンに関して基底状態の探索を試み

る。状態生成において、Bell 状態、GHZ 状態を生成する回路は既に知られている。また基底状態探索に関しても、水素分子、水素化ヘリウムイオンという分子は基底状態が既に知られている。これらの既に答えがわかっている問題に対して、提案する QOMDP プランニングアルゴリズムが解を見つけられるかどうか確認し、QOMDP の量子回路設計への適用可能性を検証する。

3. ソフトウェア開発内容

開発したソフトウェアについて解説する。全体像は図 1 に示す通りである。

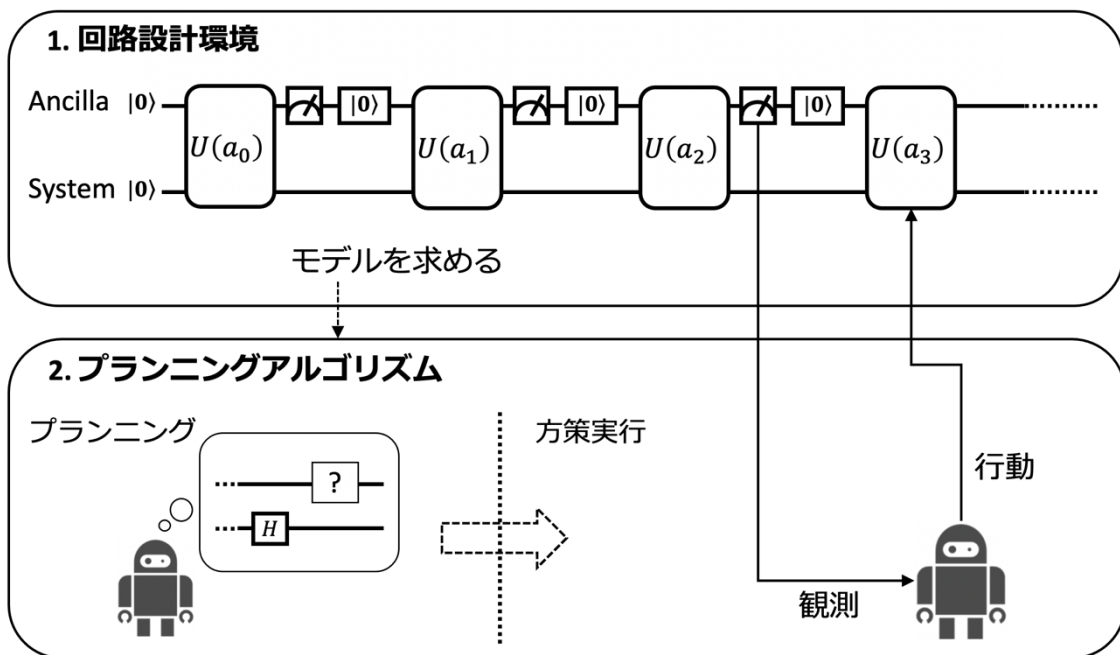


図 1. 開発プログラムの全体像

開発プログラムは、回路設計を行う環境と、環境とインタラクションしながら回路設計を行うアルゴリズムからなる。以下でそれぞれのプログラムについて詳しく説明する。

3.1 回路設計を行う環境

本プロジェクトでは部分観測を考慮した回路設計を強化学習のフレームワークで解くことを考えるので、回路において「システム」、「アンシラ」という二つの系を用意する。システムは対象としている系であり、アンシラが測定により情報を引き出すための系である。システムは測定せずアンシラのみ測定からシステムの状態を推論していくという部分観測の問題となっている。

環境の各要素は以下のように定義される。

- 状態: システムの状態

- 行動: 回路に適用するゲート
- 観測: アンシラの測定結果
- 報酬: 解くタスクごとに定義される

次に、環境の時間発展について述べる。(1)実行された行動に対応するゲートを回路に適用し、アンシラを測定し測定結果を得る。(2)システムの状態から計算される報酬を計算し、測定結果と報酬をエージェントにフィードバックする。(3)フィードバックされた情報をもとにエージェントが次の行動を実行し、(1)に戻る。上記の流れを繰り返していき、タスクの評価指標が閾値に到達するか、ステップ数が最大ステップ数に到達したらエピソードを終了する。

3.2 回路設計問題を解くプランニングアルゴリズム

本小節では、実装した回路設計アルゴリズムについて説明する。開発したアルゴリズムは、上述の回路設計環境とインタラクションをしながら、回路設計問題を具体的に解いていくアルゴリズムであり、機能はプランニングと方策実行の二つがある。以下でそれぞれについて説明する。

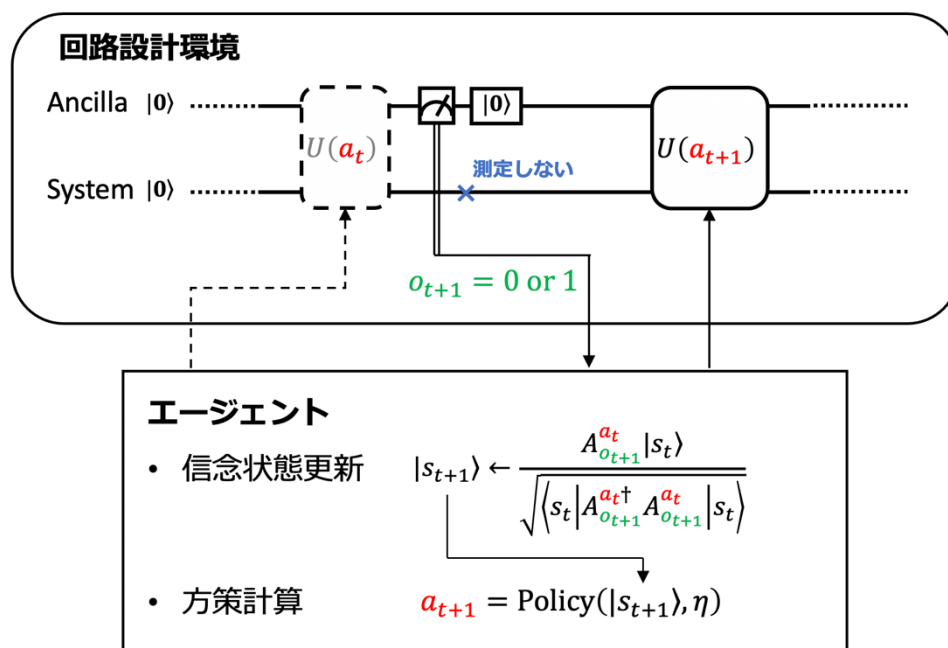


図 2. 方策実行により回路設計を行う様子

3.2.1 プランニング

プランニングは、環境とインタラクションせずに、環境のモデルを用いて方策を求めることを行う。本プロジェクトでは、方策を求める上で価値関数の更新を行い、得られた価値関数を用いて方策を計算するという方法をとる。

価値関数を更新する方法として、厳密法と近似法の二つの手法を開発した。厳密法

は、ベルマン方程式に基づき厳密な更新を行なっていく手法であり、計算量が観測数に対して指数的に増えてしまうという問題がある。そこで本プロジェクトでは、状態空間を近似し、近似的に価値関数を更新していくことで、計算量を多項式オーダーに削減する手法として近似法を開発した。

3.2.2 方策実行

方策は、上述のプランニングにより求めた価値関数を最大とする行動を求めることで計算される。この方策を用いて、エージェントが回路を設計していく方法を図2に示した。エージェントは実行した行動と、アンシラを測定して得られた観測から信念状態を更新することで、システムの状態を推論する。方策を用いて、推論した状態における行動を求め実行することで、回路にゲートをかけ設計を行なっていく。

4. 新規性・優位性

本プロジェクトで開発した量子回路設計手法は、強化学習を用いて回路設計を行なっていく手法であるが、従来の強化学習を用いた回路設計手法に対して、部分観測を導入しているという新規性がある。部分的に得られる観測からシステムの状態を推論するという形式をとることで、完全観測で定式化している従来手法に対して、圧倒的に少ない測定回数で済むことになる。また、完全観測で定式化を行うと測定により回路が終了してしまい、行動を実行するたびに始めから回路を実行しなければならないが、開発手法では、部分観測を導入することにより、補助系しか測定せず、対象系を測定しないので、回路を続けて実行していくことができるという優位性がある。

5. 期待されるユーザー価値と社会へのインパクト

開発した手法である部分観測強化学習による量子回路設計手法は、報酬の設定の仕方次第で様々な回路設計問題を解くことができるという利点がある。本プロジェクトでは具体例として状態生成、基底状態探索という二つの例を挙げたが、報酬を別のものにするすることで、別の回路設計問題にアプローチすることができるという汎用性がポイントである。本プロジェクトで開発したプログラムにおいて、ユーザーが解きたい回路設計問題の評価指標を報酬として設定し、プログラムを実行することで、回路設計を行うことができる。

6. 氏名（所属）

木村 友彰(電気通信大学大学院)