

量子アニーリングにおける自動チューニングの適用について

Adaptation of Auto-tuning for Quantum Annealing

片桐孝洋¹⁾, 森下誠²⁾, 河合直聰³⁾, 星野哲也⁴⁾, 永井亨⁵⁾

Takahiro Katagiri, Makoto Morishita, Masatoshi Kawai, Tetsuya Hosino, Toru Nagai

- 1) 博(理) 名古屋大学情報基盤センター 教授 (〒464-8601 愛知県名古屋市千種区不老町 名古屋大学情報基盤センター, E-mail: katagiri@cc.nagoya-u.ac.jp)
- 2) 修(情報) 名古屋大学 大学院情報学研究科 (〒464-8601 愛知県名古屋市千種区不老町 情報学研究科, E-mail: morishita@hpc.itc.nagoya-u.ac.jp)
- 3) 博(情報) 名古屋大学情報基盤センター 特任助教 (〒464-8601 愛知県名古屋市千種区不老町 名古屋大学情報基盤センター, E-mail: kawai@cc.nagoya-u.ac.jp)
- 4) 博(理) 名古屋大学情報基盤センター 准教授 (〒464-8601 愛知県名古屋市千種区不老町 名古屋大学情報基盤センター, E-mail: hosino@cc.nagoya-u.ac.jp)
- 5) 博(理) 名古屋大学情報基盤センター 助教 (〒464-8601 愛知県名古屋市千種区不老町 名古屋大学情報基盤センター, E-mail: nagai@cc.nagoya-u.ac.jp)

In this presentation, we will introduce the necessity and examples of applying auto-tuning (AT) in “quantum technologies,” including quantum annealers and quantum-inspired annealers, from the perspective of high-performance computing. Specifically, we will delve into the significance of employing AT technology, which becomes essential when addressing combinatorial optimization problems such as minimum vertex coverage program and support vector machines (SVM) using quantum annealer technology.

Key Words : Quantum-inspired Annealers, Auto-tuning

1. はじめに

Googleや中国科学技術大学により量子超越性を確認する実験がなされたという主張がある[1]。一方、同様の計算をアルゴリズム改良で高速化する、もしくは、最新のスーパーコンピュータ向きのアルゴリズムに変更し高性能実装を行うと、同程度の時間、もしくはより短い時間で求められる例がすぐ提示されており、現在、本当に量子超越性が達成されているかは明らかになっていない。しかしながら、量子コンピュータが注目を集めている事実はあるといえよう。

この一方で、量子コンピュータにインスピレーションを受けたアルゴリズムやハードウェアの開発が盛んにされている。これらは量子特性活用していないので、量子コンピュータではない。しかし量子「関連」技術と呼ばれており、注目がされている。特に日本においては、量子アニーラにインスピレーションを受けたハードウェアや計算方式の研究開発が盛んである。これらは、「疑似」量子アニーラと呼ばれる。疑似量子アニーラは、CMOSアニーラ（日立）[2]、デジタルアニーラ（富士通）、シミュレーテッド分岐マシン（東芝）などが知られている。

本研究ではCMOSアニーラに焦点を当て、その性能評価を目的に、性能向上のためのハイパーパラメタチューニングを対象にする。このハイパーパラメタチューニング

は、現在、AIのモデルの性能向上に必須の処理となっている。しかし歴史的には、数値計算関連のアルゴリズムや実装の性能向上のために行われているものが先である。加えて、高性能計算（HPC）分野の根幹的なテーマの1つとして知られている。我々はこの性能パラメタのチューニングに関し、ソフトウェア自動チューニング[3]（以降、ATと呼ぶ）という研究領域を立ち上げ、2000年頃から我が国が諸外国に先駆け研究活動している分野の1つである。

本稿では、このAT技術を量子関連技術に広げる研究として、疑似量子アニーラを用いた典型的な組合せ最適化問題の求解に活用する事例を紹介する。

2. CMOSアニーラ

(1) 概要

CMOSアニーラ[2]は日立製作所が開発した疑似量子アニーラである。イジングモデルの基底状態探索を行う専用ハードウェアを利用、もしくは、その原理を活用した計算を行うクラウドサービスにより利用できる。クラウドでは、「Annealing Cloud Web」[4]（GPU版、32bit, Float）が提供されている。

(2) 求解の流れ

CMOSアニーラに限らず量子アニーラにおいても、求解

の流れは以下のようになる。

- 手順1：対象となる最適化問題を設定する
- 手順2：最適化問題をイジングモデルに定式化する
- 手順3：疑似量子アニーラで問題求解を行う
- 手順4：イジングモデル上のスピン情報を取得することで、解を得る

手順3において量子アニーラでは、QUBO (Quadratic Unconstrained Binary Optimization) (二次制約なし二値最適化) による二値変数の最適化問題としてコスト関数を定義することで、最適化問題の求解が可能となる。

3. 量子アニーラにおけるATの適用

(1) 自動チューニングの流れ

ここでは、疑似量子アニーラにおけるAT適用の流れについて説明する。図-1に、ATの流れを示す。なお、本研究が対象とするのは疑似量子アニーラであるが、ハイパーパラメタチューニングの観点では疑似量子アニーラも量子アニーラも同等に扱える点に注意する。



図-1 (疑似) 量子アニーラにおけるAT

図-1では、ユーザはまず、1. 最適化問題の設定を行い、次にその最適化問題を2. QUBO定式化を行う必要がある。その後、ユーザが求める条件（解の精度、実行時間等）を満たすように、ATを行うフェーズに移る。

ここでATを行うフェーズは、以下の3つの手順を想定する：3. テスト問題の設定；4. ハイパーパラメタ設定；5. (疑似) 量子アニーラ実行。なお、手順3～5の実行は、ユーザの要求を満たすまで反復される。この反復の間、ハイパーパラメタを変化させて実行することで、チューニングを自動に行う。別の見方をすると、最適なパラメタ探索が行われているとみなすことができる。一方、機械学習の観点では、テスト問題を自動設定してチューニングをするフェーズでもあるため、強化学習をするフェーズともとらえることができる。

以上のATを行うフェーズが終了すると、ユーザの最適化条件に見合う最適なハイパーパラメタを得ることができる。最後に、最適なハイパーパラメタを用いて、(疑似) 量子アニーラで本計算（大規模問題の実行）を行

う。

以上が、本研究で実現を目指す、量子アニーラにおけるATの概要である。

(2) AT実現のための問題

図-1のATを実現するには、いくつか検討すべき問題がある。それを以下に示す。

● ATのタイミングの問題

まずATで規定されているタイミングは、以下の3つがある[5]：(I) インストール時AT；(II) 実行起動前時AT；(III) 実行時AT。この3つのタイミングは、量子アニーラにおけるAT実行でも適用できるが、いくつか考慮すべき問題がある。

(I) インストール時AT：ソフトウェアをインストール時に行うATである。この場合、3のテスト問題の設定には、ユーザ実行する際の問題の詳細がわかつていないため、ユーザ問題の特性が予測可能な問題設定による実行、および、適切な問題を用いてユーザ問題特性に影響しないが性能に影響するハイパーパラメタのチューニングをする。加えて、実行時にユーザの問題が確定した段階で、インストール時のチューニング情報をもとに、最適なハイパーパラメタを予想する予測モデルの生成とセットで行われるAT方式である。

(II) 実行起動前時AT：ユーザが実行する問題を確定した段階で、ユーザがAT実行を指示して行うATである。

(I) インストール時ATと異なり、ユーザの問題特性の情報が入手可能であり、精度の高いチューニングが可能である。一方、大規模な問題を指定して実行時起動前ATを行うと、AT時間の増大につながり、ATコストの増大を招く。主に、ユーザが指定する同一の問題を何回も解く場合に使われるAT方式である。

(III) 実行時AT：対象の手続き、関数もしくはライブラリが実行されるときに行うATである。ユーザがAT実行の指示を与えなくてもATを実行する。ユーザが指定する問題特性を必ず入手できるため、精度の高いATが行うことができるが、実行ごとにATをするためAT実行のコストが高い。そのため、総合的な実行時間の観点で、ATをしない場合に対してメリットが得られない場合がある。

● ATのためのテスト問題の設定の問題

ATを行うためには、何らかのテスト問題を指定しなくてはならない。この場合のテスト問題指定は、ATのタイミングにより異なる。

(I) インストール時AT：テスト問題指定は、ソフトウェア開発者が行う。ユーザが指定する問題が不明なため、実行されそうな問題を事前に集めて学習し、何らかの学習モデルを構築する必要がある。

(II) 実行起動前時AT：ユーザが実行する問題を直接与える。ATに必要なテスト問題の設定は、以下の2つが考えらえる：(A) ユーザが与えた問題そのものを利用する；(B) ユーザが与えた問題を、何らかの手法で縮小した問題を自動生成して利用する；

(B) の方法で、AT時間の削減を図ることができるが、ATによる解の精度は低下する。また、なんらかの問題縮小の方法は、行列計算では、元の行列要素をランダムサンプリングにより抽出して小規模行列を作成することが多い（ランダマイズドアルゴリズムと呼ぶ）。ここで、量子アニーラでは組合せ問題を対象にするため、グラフ問題が対象になる。しかし、グラフ問題をランダムサンプリングにより問題縮小したテスト問題の活用によるATの研究はほとんど行われておらず、重要な今後の課題であると考えている。

(III) 実行時AT：ユーザが与えた問題を必ず設定できる。

(3) 量子アニーラにおけるハイパーパラメタ実例

ここでは、疑似量子アニーラにおけるハイパーパラメタの実例を表-1に示す。

表-1 疑似量子アニーラにおけるハイパーパラメタ

名称	概要
Wa	QUBO制約項の重み
Wb	QUBOコスト項の重み
chain_strength	エッジの重み
temperature_num_steps	アニーリングのステップ数
temperature_step_length	アニーリングのステップ長
temperature_initial	アニーリングの初期温度
temperature_target	アニーリングの最終温度

表-1では、QUBOに起因するハイパーパラメタと、アニーリングに起因するハイパーパラメタがある。なお、QUBOに起因するハイパーパラメタは、疑似量子アニーラに限定しない、量子アニーラにおける汎用的なハイパーパラメタである。加えて、このハイパーパラメタは解の精度に大きく影響することが知られており、量子アニーラにおいて重要な性能パラメタである。現状、このWaとWbパラメタについては、適当な値が設定するか、研究者がいくつか実行してよさそうな値が設定されていることが多いと予想される。そのため、性能チューニングの研究においては大問題と考えている。

加えて、表-1に示していないハイパーパラメタも量子アニーラでは存在する。その多くは、QUBO定式化で現れる性能パラメタである。この多くも、解の精度に影響するため、量子アニーラ分野における重要なハイパーパラメタになる。

(4) QUBO実例

ここでは量子アニーラにおけるQUBO記載の実例を紹介する。

● 最小頂点被覆問題

最小頂点被覆問題は、組合せ最適化問題として典型的な問題として知られており、NP困難な問題である。グラフ $G(V, E)$ において、それぞれの枝 e について、端点のいずれか少なくとも一方が V' に含まれるような頂点集合 V の部分集合 V' のうち、頂点集合の数 $|V'|$ が最小になるものを求める問題である。

最小頂点被覆問題のQUBOは、式(1)のようになる。

$$H = W_a \sum_{(u,v) \in E} (1 - x_u)(1 - x_v) + W_b \sum_{v \in V'} x_v \quad (1)$$

ここで、 $x_u, x_v \in \{0,1\}$ である。

式(1)は、ある種のエネルギーを定義しており、小さいほど最適である。式(1)の前半の項は、制約項と呼ばれる。この最小頂点被覆問題の場合、被覆される頂点であると0となり最小となるように設定されている。また式(1)の後半の項は、コスト項と呼ばれる。コスト項は、低ければ低いほど評価値が高いように設定される。この最小頂点被覆問題では、被覆する頂点数が少ないほど良い解なので、被覆している頂点数をコスト項に設定している。

ここで、式(1)においては、制約項の重み W_a 、コスト項の重み W_b がある点に注意する。この重みは、制約項とコスト項どちらを重視するかのパラメタである。このパラメタの最適値は、与える問題の性質により異なることから、解の精度に関わる重要なハイパーパラメタになる。そのため、ハイパーパラメタチューニングが原理的に必須となる。

4. 予備評価

(1) 実験環境

本予備実験では、CMOSアニーリングマシンとして、Annealing Cloud Web API V2 (GPU version 32bit (float)) [4] を利用した。端末となるPCは、MacBookAir (macOS Big Sur) であり、1.6GHz Dual Core Intel Core i5 (メモリ8GB) である。プログラミングは、Python (Version 3.8.2)を利用している。また、Cloud WEB上の実行には、FIXSTARS Amplify [6]を利用している。これは、CMOSアニーラを使うためのWeb APIであり、本実験でのバージョンは0.5.13である。

(2) 問題設定

本予備実験では、最小頂点被覆問題の解の品質を評価可能な人工問題を設定して、CMOSアニーラによる解の品質チェックを行った。具体的には、2次元格子状のグラフ問題をベンチマークとして与える。図-2に例を示す。

図-2から、最適解の形状は自明なため、CMOSアニーラが最適解を求めているかのチェックが可能である。

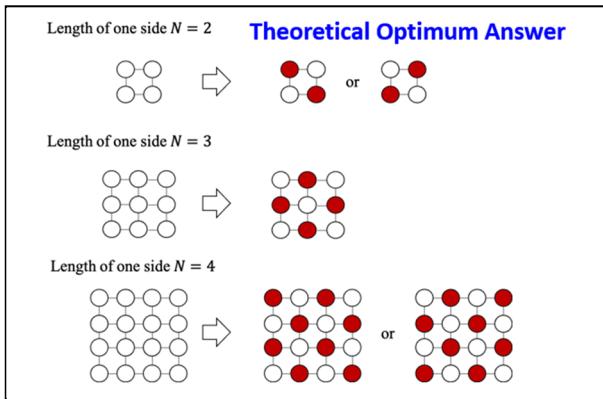


図-2 最小頂点被覆問題のベンチマーク

(3) 実験結果

ここでは、式(1)の W_a , W_b , 及び表-1 の $chain_strength$ に1.00を与えたデフォルト実行の最適解求解の精度（100回の実行に対して、最適解を出す割合[%]）と、ハイパーパラメタチューニングにATを行った場合の最適解求解の精度を評価した。なお、ATソフトウェアとして、Optuna[7]を利用している。

表-2にATの効果を示す。

表-2 自動チューニングの効果

(a) デフォルト実行

問題 サイズ	W_a	W_b	$chain_strength$	最適解 回答率
$N=3$	1.00	1.00	1.00	5%

(b) 自動チューニング適用後

問題 サイズ	W_a	W_b	$chain_strength$	最適解 回答率
$N=3$	9.76	0.01	6.42	86%

表-2から、デフォルト実行では、最適解回答率は5%と極めて低いが、ATを適用すると86%と劇的に最適解を回答する率が向上している。このように、疑似量子アニーラにおいてもハイパーパラメタチューニングは極めて重要である。

量子アニーラにおけるハイパーパラメタチューニングは従来のAT技術が適用できることは間違いないが、量子アニーラ特有のAT上の問題は明らかになっていない。そのため我々は、多くの事例によるATの効果検証が必須であると考えている。

5. おわりに

本研究では、量子アニーラ、および疑似量子アニーラを含む量子関連技術へのATへの適用について述べた。量子アニーラにおいては、QUBOへの定式化が必須であるが、その際に解の精度に影響するハイパーパラメタが原理上、必ず現れる。そのため、ハイパーパラメタチューニングが

必須となるが、従来はその性能チューニングについて自動化やコスト削減、もしくは十分なチューニングが試みられていないと予想している。

この一方で我々は2000年ごろから、数値計算に現れる性能パラメタの自動チューニング問題に取り組んできた。そのため、パラメタチューニングのノウハウが多く蓄積されており、AT技術を量子関連研究へ展開することは、今後の量子コンピューティングの発展に大きく資するものであると信じている。

量子関連技術へのATの適用は、まだ始まったばかりである。量子アニーラに加えて、量子回路型の量子コンピュータへのAT技術の展開も期待される。加えて、量子回路シミュレータの高速化へのAT技術の展開は既に始められている[8]。量子コンピュータ技術へのATの適用は世界的に見てもほとんど行われてない。そのため、ATの観点での問題も明らかになっていないため、多くの今後の課題が山積している。

謝辞: 本研究は、学際大規模情報基盤共同利用・共同研究拠点、および、革新的ハイパフォーマンス・コンピューティング・インフラ（課題番号：jh230005）の支援による。CMOS アニーラの利用に関して助言を頂いた、日立製作所の小塙和夫氏と山岡雅直氏に感謝します。また、Amplify の利用に関して助言を頂いた、株式会社フィックスターズの松田佳希氏に感謝します。

参考文献

- [1] Gibney, E.: Hello quantum world! Google publishes landmark quantum supremacy claim, *Nature*, Vol. 574, pp. 461-462, 2019.
- [2] Yamaoka, M. et al.: 20k-spin Ising chip for combinational optimization problem with CMOS annealing, *Proceedings of 2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*, 2015.
- [3] Katagiri, T. and Takahashi, D.: Japanese Autotuning Research: Autotuning Languages and FFT, *Proceedings of the IEEE*, Vol.106, Issue 11, pp. 2056 – 2067, 2018.
- [4] Annealing Cloud Web,
<https://annealing-cloud.com/ja/index.html>
(2024年3月30日閲覧)
- [5] Katagiri, T. et al.: FIBER: A generalized framework for auto-tuning software, *Proceedings of High Performance Computing: 5th International Symposium*, ISHPC 2003, LNCS, Vol. 2858 pp. 146-159, 2003.
- [6] FIXSTARS Amplify, 量子コンピューティングプラットフォーム,
FIXSTARS社, <https://amplify.fixstars.com/ja/>
(2024年3月30日閲覧)
- [7] Preferred Networks, Inc. : An open source hyperparameter optimization framework to automate hyperparameter search. <https://optuna.org/> (2024年3月30日閲覧)

- [8] 森下誠ほか：量子コンピューティングへの自動チューニングの適用と評価，*研究報告ハイパフォーマン*

スコンピューティング(HPC), Vol. 2023-HPC-188, No. 2, pp. 1 – 7, 2023.