

読んだ論文の共有

tax_free

東京工業大学 情報理工学院 数理・計算科学系 学士課程 3 年

May 7, 2024

- 1 Quantum annealing-driven branch and bound for the single machine total weighted number of tardy jobs scheduling problem [1]
- 2 Balanced Allocation of Educational Resources Based on Parallel Genetic Algorithm [2]
- 3 Large Language Models as Optimizers [3]

Quantum annealing-driven branch and bound for the single machine total weighted number of tardy jobs scheduling problem [1]

論文情報

著者	Bozejko, Wojciech and Pempera, Jaroslaw and Uchonski, Mariusz and Wodecki, Mieczyslaw
雑誌	FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE 2024, 155, 245-255.
url	https://www.sciencedirect.com/science/article/pii/S0167739X2400061X?via%3Dihub

モチベーション

量子アニーリングに興味があり、実行原理などの低レイヤーを扱っている論文ではなく、応用寄りの論文を読みかったから

どんなもの?

どんなもの

- 分枝限定法のアルゴリズムに量子アニーリングメタヒューリスティックを組み込んで目的関数の上界と下界を計算できるようにした
- 量子コンピュータからの結果は確率的であるが、CPU と QPU を交互に使用するハイブリッドアルゴリズムによって最適解を保証する

分枝限定法とは

組合せ最適化問題を解くための最適化アルゴリズムの 1 つで、問題を小さな部分問題に分割し各部分問題を再帰的に解くことで最適解を探索する。各部分問題の上界と下界を計算することで、最適解を含まない部分問題を探索範囲から落とす (枝刈り) を行って効率的に探索する。

先行研究と比べてどこがすごい？

title

- 量子アニーリングを用いることで計算速度を向上させた点
- 従来の量子アルゴリズムでは確率的な振る舞いによって安定性が低かったが、これを小さくして安定性を改善している点
- 量子アルゴリズム × 古典アルゴリズムのハイブリッドアルゴリズムの可能性を示している点

技術や手法のキモはどこ？

老子アニーリングの利用

- QPU を用いてラグランジュ緩和法に基づいて目的関数の下界を与える．これによって，既存の手法では扱うことが難しい領域を効率的に探索できる
- QPU が扱えるように問題を Binary Quadratic Programming に変換している

分枝限定法との統合

量子アニーリングを分枝限定法を組み合わせることで，上界と下界を動的に計算し，より迅速に最適化を探索することができる．

ラグランジュ緩和法とは

制約付き最適化問題を解くための数学的アプローチの一つで，問題の制約を目的関数に組み込むことで，元の問題をより簡単な問題に変換する．このとき，各制約に非負のラグランジュ乗数を割り当て，新しい目的関数を最小化する．

どうやって有効だと検証した?

単一機械スケジューリング問題とは

- 単一の機械で n 個のジョブを実行することを考える
- 機械は 1 つのジョブだけを実行することができ、その上でジョブの実行順序を最適化し、各ジョブに紐付けられた遅延ジョブの重み付き数の和を最小化する
- NP-Hard であることが知られている
- 検証ではスケジューリング問題のベンチマークセットを使用いた

結果

- 早い段階で最適解を発見できた
- アルゴリズムが計算した上界と下界の差が非常に小さく、正確に見積もれていることがわかった
- 古典コンピュータのみでの実行と比較して計算時間が大幅に短縮された

議論はある？

古典アルゴリズムとの比較

- 量子アルゴリズムが古典アルゴリズムに対してどれほど優位性があるかについては、より詳細な比較を行う必要がある
- 特に大規模な問題での比較は数が少ない

アルゴリズムの適応範囲

- 提案された手法が他の NP-Hard に対してどの程度応用可能かについて検証する必要がある

→ 現状の量子アニーラの限界もあるので、どの程度まで実問題に応用可能かを調べることは重要であるが、こういったベンチマークがあるかが分かっていない

Balanced Allocation of Educational Resources Based on Parallel Genetic Algorithm [2]

論文情報

著者	Jan, Naeem and Qiu, Ming
雑誌	Mathematical Problems in Engineering, 2022, 7517267
url	https://www.hindawi.com/journals/mpe/2022/7517267/

モチベーション

教育格差の是正に関連する研究を探していたところ、この論文が教育資源の均等な配分に関する研究であることを知り、興味を持ったから。

どんなもの？

どんなもの？

- 限られた教育リソースを効果的に配置するために、教育のコストに対するパフォーマンスを定式化し、それを GA と PSO を用いて最適化した
- 検証には山西省の大学のデータを利用して、モデルをシミュレートし、指標の向上が見られた

GA とは

遺伝的アルゴリズム (Genetic Algorithm) の略で、生物の進化のメカニズムを模倣した確率的な最適化手法。問題の解を個体として表現し、複数の個体からなる集団を進化させることで最適解を探索する。

PSO とは

粒子群最適化 (Particle Swarm Optimization) の略で、鳥や魚の群れの社会的な行動を模倣した確率的な最適化手法。多数の particle を用いて解空間を探索し、各 particle が最適解を見つけるために強調しながら動作する。

先行研究と比べてどこがすごい？

先行研究と比べてどこがすごい？

- GA と PSO を組み合わせて使用してる点
- 単一目的の最適化ではなく複数目的 (利用効率と配分効率) の最適化をしている点
- 独自の新しい評価システムを導入している点 (以下の画像)
- 理論だけではなく実際のデータを用いて検証している点

TABLE 1: Education resource input-output index system in colleges and universities.

First-level index dimension	Secondary factor index	Tertiary measurement index	Symbol
Educational resources input indicators	Human resources	Number of full-time and part-time teachers	I_1
		Administration number of teachers	I_2
		The number of outside school teachers with corporate backgrounds	I_3
	Material resources	The total value of teaching equipment	I_4
		The floor area of practice platform	I_5
		Area of education base	I_6
	Financial resources	The special fund's investment	I_7
Educational resource output indicators	Cultivation of talents	Number of students in education	O_1
		Student enterprises	O_2
		Awarded in discipline competitions at or above the provincial level	O_3
	Scientific research	Number of academic papers, works, and topics	O_4
	Social services	The output value of scientific and technological achievements	O_5

技術や手法のキモはどこ？

PSO の定式化

- より厳密には PSO の multi objective version である MOPSO を用いている
- 教師の数や設備投資の量といった教育リソースの配分の具体的なパラメータを Position としている
- 更新式は一般的な PSO に従っていて、(おそらく) 乱数を用いて少し外乱を加えている

GA の定式化

- PSO の Position と同じようにして Chromosome を定義している
- 通常の GA と同様に、選択、交叉、突然変異を経て解が更新される

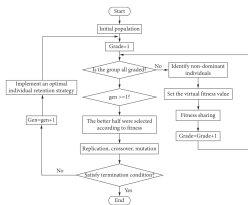


FIGURE 1: Flow chart of multiobjective genetic algorithm.

どうやって有効だと検証した?

検証データについて

- 山西省に位置する大学からの実際のデータを使用
- 2017 年から 2019 年までの革新と起業教育に関連する教育リソースの統計データが含まれる
- これらのデータは次元や単位の違いを標準化するために量化および正規化された

結果について

- 利用効率: 0.695 から 1.010 の範囲だった値が 1.026 から 1.084 の範囲に改善した (平均で 18.72 % の向上)
 - 配分効率: 0.123 から 0.187 の範囲だった値が, 平均で 0.189 に改善した (平均で 20.89 % の向上)
- 教育リソースがより効果的に使用され, かつリソースの分配の公平性が向上した

議論はある？

効果の検証について

- 数値上は教育の質や平等性が向上したが、これらが本当に意味があるかはより長期間・詳細なデータを用いた検証が必要である

アルゴリズムの適応条件について

- 今回の問題は上手く解くことができたが、他の教育リソースに関する問題や異なる分配問題に応用できるかは未知数である

→ 日本の教育環境への応用やより一般的なパッケージにして多くの関係者に使ってもらえるようにするなどの展開が考えられる

Large Language Models as Optimizers [3]

論文情報

著者	Chengrun Yang and Xuezhi Wang and Yifeng Lu and Hanxiao Liu and Quoc V. Le and Denny Zhou and Xinyun Chen
雑誌	arXiv
url	https://arxiv.org/abs/2309.03409

モチベーション

LLM を用いて最適化問題を解くことに関心があったから

どんなもの? & 先行研究と比べてどこがすごい?

どんなもの?

- 自然言語で記述された最適化タスクを LLM を用いて解を生成するアプローチ Optimization by PROMpting (OPRO) を提案した.
- OPRO を用いて数理最適化の問題やプロンプト最適化の問題で性能を検証した.

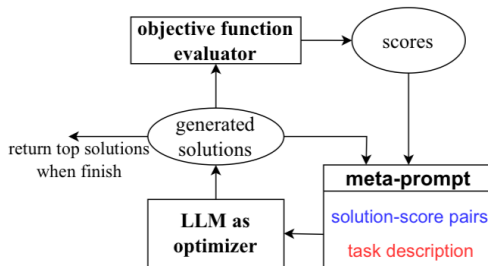
先行研究と比べてどこがすごい?

- LLM を最適化の solver として利用していること
- プロンプト最適化を自動化していること
- 最適化したプロンプトが同じドメインの他の問題でも使用できること
- 複数の LLM を用いて性能を評価していること

技術や手法のキモはどこ？

技術的なキモ

- 数理最適化用にメタプロンプトを設計していること
- 入力を自然言語として数理最適化の問題を解けること (解けるとは言っていない)
- プロンプト最適化で性能を向上させることができることを示し、また同じドメインのタスクでも使用できることを示したこと



どうやって有効だと検証した?

数理最適化について

- 線型回帰: ブラックボックス最適化として機能しているが、十分ではない
- TSP: $n=10$ までは最適解を得るが、それより大きい n ではヒューリスティックと比較して大幅に性能が下がる

→ 小さい n では大域的な最適解を得られるが、大きい n ではまだまだ

プロンプト最適化について

タスクの内容: あるタスクの精度を最大化するようなプロンプトを見つける

- 複数のベンチマーク (GSM8K や BBH など) で人間が作成したプロンプトを大きく超えるスコアを得た
- 具体的には GSM8K では最大で 8% 程度, BBH では最大で 50% 程度の向上が見られた

→ 大きく改善した結果を得られた

議論はある？




数理最適化について

- 規模が大きくなると性能が大幅に低下してしまうので改善が必要がある
- 勾配に関する情報を用いることができれば，より性能が向上することが見込まれる

プロンプト最適化について

- 初期プロンプトに大きく影響されるので，収束をコントロールできるようになる必要がある
- 検証データを用いていないので過学習になっている可能性があるので，過学習を抑える方法を考える必要がある

→ LLM 単体で数理最適化用にチューニングされたソルバーには勝てないと思うので，より使いやすいインターフェースとして使用する方がよさそう

-  Bozejko, Wojciech and Pempera, Jaroslaw and Uchonski, Mariusz and Wodecki, Mieczyslaw *Quantum annealing-driven branch and bound for the single machine total weighted number of tardy jobs scheduling problem*. FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE 2024, 155, 245-255.
-  Jan, Naeem and Qiu, Ming *Balanced Allocation of Educational Resources Based on Parallel Genetic Algorithm*. Mathematical Problems in Engineering, 2022, 7517267
-  Chengrun Yang and Xuezhi Wang and Yifeng Lu and Hanxiao Liu and Quoc V. Le and Denny Zhou and Xinyun Chen *Large Language Models as Optimizers*. arXiv, 2023.