

第4回 シンポジウム「富岳百景」スーパーコンピュータ「富岳」シンポジウム 質問回答一覧

2024年12月25日開催シンポジウムの講演内容へのご質問をありがとうございました

*本一覧は2024年12月25日～2025年1月6日の質問受付期間に、
質問フォーム・zoomチャット(配信当日のみ)・SNS(X)投稿(指定タグ #富岳百景シンポ 24)からお寄せいただいた質問の中から
一部を選定し回答を公開するものです。

= 質問と回答 =

講演1: 「富岳」を用いた大規模な AI の事前学習	1
講演2: 次世代気象レーダによる未来のゲリラ豪雨予測 ～大阪・関西万博におけるリアルタイム配信の実証～	1
講演4: シミュレーションと AI で解明する宇宙の構造と進化	2
パネルディスカッション: 科学は AI でどう進化する? ～ AI for Science ～	3～5

セッション1 政策対応利用課題

講演1: 「富岳」を用いた大規模な AI の事前学習

(東京科学大学 総合研究院 スーパーコンピューティング研究センター 教授 横田 理央 氏)

Q1 質問:

「富岳」で LLM を訓練した目的は何だったのでしょうか?

最初の説明にもあった通り、CPU のみの「富岳」は他の GPU のあるスーパーコンピュータと比較して占有するプロセス数の割には訓練にかかる時間が多いと感じています。

また、「富岳」で LLM を訓練するにあたって、様々なチューニングに関する知見があったことを学びましたが、「富岳」以外で応用可能なのでしょうか?

A1 回答:

CPU は LLM に向いていない」という定性的な知見だけでは「富岳」の次の次世代計算基盤の設計の指針にはなりません。具体的にどのくらいの演算性能、メモリバンド幅、I/O バンド幅、ノード間バンド幅・レイテンシがあれば、LLM の学習がどの程度速くなるのかを定量的に見積もる必要があります。また、当時は GPU 資源が国内にそんなになかったため、数万ノードを何ヶ月も占有できるマシンは「富岳」だけでした。GPU がいくら速くても少量・短期間しか使えないのでは LLM は学習できません。「富岳」で得られた通信の高速化の知見はアルゴリズムレベルでの知見なので違うハードウェアでも有効です。

(回答者: 東京科学大学 総合研究院 スーパーコンピューティング研究センター 教授 横田 理央 氏)

セッション1 政策対応利用課題

講演2: 次世代気象レーダによる未来のゲリラ豪雨予測 ～大阪・関西万博におけるリアルタイム配信の実証～

(総務省 国際戦略局 技術政策課研究推進室 富士原 大介 氏)

Q1 質問:

「富岳」を使って、30分先までの予測が可能であるというお話がありましたが、実際にデータを取得してスーパーコンピュータを使って計算し、計算結果を得るまでにどれくらい時間がかかるものなのでしょうか?

リアルタイム運用にあたってどんなボトルネックや、障壁があるのでしょうか?

A1 回答:

リアルタイムといっても、莫大な計算を行うので、過去の実績では、計算結果を得るまで、3分程度かかっています。

今回も同様の時間で計算を実施し、そのデータを配信することを目指しています。

MP-PAWR を2台使うことで、さらに計算量が増えたのですが、単純に「富岳」の専有を増やせばいいというわけではありません。

実は、「富岳」に入力するデータ量が増えると難易度が上がるため、その中で、計算範囲やアンサンブル数を調整する等の手法を工夫を行っており、予測の精度を前回よりも向上させることを目指して、理化学研究所を中心に調整や研究を行っているところです。

(回答者: 総務省 国際戦略局 技術政策課研究推進室 富士原 大介 氏)

Q2 質問 :

次世代気象レーダの必要性が理解できました。知らないことばかりで有益でした。ありがとうございました。

さて、次世代レーダ (MP-PAWR) ではデータ取得間隔が 5 分から 30 秒に短縮されるとのことでした。気象予測には、台風の進路や線状降水帯の発生予測等 更に多くの予測があるかと思います。「富岳」のようなスパコンを必要とするのはどの予測になるでしょうか。

また、それらを高精度に予測しようとするれば、データ取得間隔の更なる短縮が必要でしょうか。

A2 回答 :

講演でもご紹介した、数値予報と呼ばれる予測手法を行う場合には、スーパーコンピュータが必要で、数値予報によって、普段の天気予報はもちろん、台風の進路や、線状降水帯の発生予測に活用されています。気象庁でも、スーパーコンピュータの導入を行い、予測精度の向上を図っています。

特に、「富岳」のような計算能力が求められる場合としては、例えば今回のように従来よりも膨大なデータを使って高精度に予測する場です。

ご指摘の通り、高精度に予測するための方法として、データの取得間隔を短縮するというのが選択肢の一つです。

例えば、理化学研究所では、今回ご紹介した 30 秒ごとにデータを取得できるフェーズドアレイ気象レーダを、仮想的に配置した場合に、線状降水帯の豪雨予測精度が改善できることを示す研究を行っています。(以下参考リンクになります。)

https://www.riken.jp/press/2022/20220307_1/index.html

(回答者 : 総務省 国際戦略局 技術政策課研究推進室 富士原 大介 氏)

セッション 2 「富岳」成果創出加速プログラム

講演 4 : シミュレーションと AI で解明する宇宙の構造と進化

(筑波大学 計算科学研究センター 教授 大須賀 健 氏)

Q1 質問 :

ブラックホールに限らず、将来的に宇宙全てを人間がシミュレートすることは可能ですか？

また、宇宙全体を「富岳」でシミュレートしようとしたとき、何台必要になりますか？

A1 回答 :

要素を個別のシミュレーションで解明し、その知見を統合することで、宇宙の全てを理解することは可能と思っています。各要素を完全に理解には「富岳」をはるかに超える計算機パワーが必要ですが、最終の統合モデルは意外とシンプルで、さほど大規模な計算機が必要ないのではないかと考えています。あくまで私の直感です。

(回答者 : 筑波大学 計算科学研究センター 教授 大須賀 健 氏)

Q2 質問 :

ビッグバンから現在の宇宙に関係した壮大なスケールの計算があるということが驚きでした。これまでどのスパコンでも良いのですが、時間的な制約で解けなかった計算の所要時間 (予想) はどれくらいでしょうか。自分のケースでは、1 か月計算した時点で 50 年程度かかることがわかり 計算を途中で中断しました。今後、AI を活用すれば、シミュレーションに要する時間を短縮できると感じました。

A2 回答 :

宇宙におけるほぼ全ての問題は、必要な分解能が高いにもかかわらず、長時間の進化を追跡する必要があるため、まともに試算すると宇宙年齢かけても終わらない計算になることが多いです。このため、現存の計算機で最大限の成果を上げられるように、問題を切り分けたり、目標を絞り込むなど、問題を再設定することが研究者の重要な仕事の一つとなっております。

(回答者 : 筑波大学 計算科学研究センター 教授 大須賀 健 氏)

Q1 質問：

AI モデルのスケールリングにおいて、データの多様性や代表性はどのように確保されていますか？
特に、特定の言語や文化に特化したモデルの開発における課題について教えてください。

A1 回答：

学習に用いたデータの大半はインターネット中のデータを集積した CommonCrawl をもとに作成しています。多様性の確保は重複削除という操作によって類似の文書を除くことで確保しています。また、教科書などの質の高いデータも色々なところからとってきてデータに混ぜています。

英語で獲得した知識は日本語を少量入れて学習することで日本語に転移することが分かっており、日本語のデータをそこまで大量に集める必要はありませんでした。

(回答者：東京科学大学 総合研究院 スーパーコンピューティング研究センター 教授 横田 理央 氏)

Q2 質問：

物質によって最適な構造解析法が変わってくると思うのですが(NMR、X 線、中性子線、etc...)、AI によって対象の物質にはどの解析法が最適か導き出すことはできると思いますか？

A2 回答：

明らかにしたい物質のスケールによってどのような測定方法を使うべきであるのかは確立されておりますので、ChatGPT などが既に学習している可能性は高いです。そのため、まず ChatGPT 等により情報を得てから専門書または専門家に聞く方が効率よく実験計画を進められるのではないかと思います。

(回答者：関西大学 化学生命工学部 准教授 藤本 和士 氏)

Q3 質問：

「富岳」は GPU を使用せず CPU のみで構成されていますが、この設計は他の GPU ベースのスパコン(例:Summit や Aurora)と比べてどのようなメリットとデメリットがありますか？

A3 回答：

「富岳」に採用されている CPU は携帯などに入っている ARM の CPU をベースに富士通が作ったものであり、低電力消費に特化した設計になっています。これから先、ボトルネックは演算速度よりも電力消費になっていくので、CPU、GPU に限らず、このような低電力消費のプロセッサで LLM を学習できるかを試したかったというのがあります。

(回答者：東京科学大学 総合研究院 スーパーコンピューティング研究センター 教授 横田 理央 氏)

Q4 質問：

分散並列深層学習での 6 次元トポロジーとは？

A4 回答：

大規模なモデルの学習は数百 GPU を用いても何ヶ月もかかります。たくさんの GPU を用いて分散並列化を行って深層学習を行うことを分散並列深層学習と呼びます。これをやらないと何十年もかかってしまいます。

6 次元トポロジーは「富岳」のノード間の接続の仕方のことです。一直線にノードを繋げると 1 次元、平面上にノードを配置して x,y 方向に接続するのが 2 次元トポロジーです。要はノード間がどれだけ密に接続されているかを表しています。

(回答者：東京科学大学 総合研究院 スーパーコンピューティング研究センター 教授 横田 理央 氏)

Q5 質問：

シミュレーション限界の時間問題で、具体的に 1 秒未満の計算がなぜ難しいのですか？

A5 回答：

CPU のクロック数（1 秒間の演算回数）は残念ながら、私の大学時代から変わっておりません。

例えば、MD 計算では一般的に、1 回のステップ（1 回に進められる時間）で 1fs (10^{-15} 秒)しか進めません。

つまり、1 秒の MD 計算には 10^{15} 回の繰り返し計算を行わなければいけません。1 秒間の演算回数が変わっておりませんので、進められる時間は自動的に同じままとなります。

しかしながら、規模（原子数）の問題は「京」や「富岳」のおかげで解決しております。これらのスーパーコンピュータは多数の CPU があるため、大きな規模の計算が可能となり、これまで不可能であったウイルスや燃料電池触媒層の計算が可能となりました。

（回答者：関西大学 化学生命工学部 准教授 藤本 和士 氏）

Q6 質問：

先程のシミュレーション限界の時間問題(1 秒未満の計算の難しさ)でのハードウェア問題と言われましたが、例えば、一つとして、ハードウェア的にマイクロ空間での分子相互間エネルギー計算に特化したデバイスチップのようなものはどうでしょうか？

A6 回答：

そのとおりです。実際 MD 計算専用のデバイスも開発されております。

例えば、理研の研究者が開発した MDGRAPE やアメリカの Anton などがあります。

（回答者：関西大学 化学生命工学部 准教授 藤本 和士 氏）

Q7 質問：

光エネルギー分布のところ、その光エネルギーとは何を指すのですか？

A7 回答：

光の粒子である光子は、波としての性質も持っていて、波の波長に反比例するエネルギーを持っていることが知られています。

例として、青い光と赤い光を比べると、波長の短い青い光のほうが光子一個あたりのエネルギーが大きいこととなります。

（回答者：筑波大学 計算科学研究センター 教授 大須賀 健 氏）

Q8 質問：

皆さまのお話の中で、データが揃っていないものについては観測(データ集め)とシミュレーションを組み合わせで行うというお話があったと思います。

科学においてデータは多ければ多いほど良いように思いますが、宇宙の場合は地上からの観測でもロケット等を用いてデータを集める場合でも非常に 1 つのデータの価値が大きいように思いますが、そのあたりの戦略をどのように立てられているか、どのような苦労があるか教えていただきたいです。

A8 回答：

入手できる限られたデータから、新たな仮説の提案や、もしくは、既存の仮説の一部を棄却することが可能となります。

そして、有力な仮説を確かめるために最も有効な観測は何かを理論的に絞り込み、次なる観測目標を設定します。この観測から理論へ、そして理論から観測へというサイクルにより、観測と理論を最大効率で活かす研究が可能となります。

しかし、技術の限界があったり、理論研究の不定性があったりするので、理想的に進まないところが難しいところでもあり、やりがいのあるところでもあります。

（回答者：筑波大学 計算科学研究センター 教授 大須賀 健 氏）

Q9 質問：

綺麗な画像というのは、先ほど講演の中で仰っていたような解像度が高いという意味でしょうか？

A9 回答：

解像度に加え、ノイズが小さい（S/N 比が大きい）画像を指します。患者個別を対象とした解析では、医用計測データ（画像）が個々の違いを反映した重要な情報となるため、その精度を高めることが重要となります。

（回答者：東京科学大学 工学院 機械系 教授 伊井 仁志 氏）

Q10 質問：

血管内での血液の流れにおける“キャビティ流れ問題”とは、もっと細かい推定理論などがあるのですか？

A10 回答：

キャビティ流れは、くぼみ内の流体の流れを意味し、工学分野を始め自然界でも広く見られる流れです。

脳動脈瘤は血管から生じた突起物であり、その内部の流れはキャビティ流れと捉えることができます。

今回の発表で紹介した研究例の一つでは、キャビティ流れの中でも最も単純な「矩形領域において、一辺を入口として、その入口が横に滑る状況で生じる流れ」を対象に、提案した手法の妥当性を示しています。実際の動脈瘤内流れを考慮すると、くぼみ形状と入口速度がより複雑になるため、問題の難易度が上がりますが、適用する推定理論は基本的に同じものを使っています。

（回答者：東京科学大学 工学院 機械系 教授 伊井 仁志 氏）

以上