

## 審 査 の 結 果 の 要 旨

氏 名 張 金哲

本論文は、全六章からなり、第一章では、**NMR**（核磁気共鳴法）分光法による分子同定に関する、これまでの歴史と、従来手法について述べている。これまでは、**NMR** スペクトルから、サンプル中に含まれる分子を同定する作業は、専門家が **NMR** データベースなどの情報を用いて手動で行っており、効率に限界がある。近年、ロボティクスを用いた自動運転研究システムが話題となっているが、**NMR** による手動による同定は、このようなシステムにおけるボトルネックとなり、効率の向上を妨げる可能性がある。本研究では、深層学習と量子化学計算を用いて分子同定を行うアルゴリズム **NMR-TS** を提案する。このアルゴリズムは、データベースに含まれていない新規の分子を同定する能力を持つ。

第二章では、**NMR** 分光法の基礎に関して述べている。 $^1\text{H}$ 、 $^{13}\text{C}$  などの異なる分光法の種類や、化学構造式に含まれる部分構造と、化学シフトとの関連などが概観されている。手動で分子同定を行う場合には、ピーク割当を行って、スペクトルに存在するピークを分子の部分構造に対応づけた上で、それらをパズルのように纏め上げることで、分子を推定する。データベースに存在しない、未知の分子に関して同定を行うには、深い化学に関する知識と熟練が必要である。

第三章では、**NMR-TS** の基礎となる新規分子生成アルゴリズムに関しての従来研究を概観している。特に、モンテカルロ木探索と **Recurrent Neural Network(RNN)**の組み合わせによって分子を生成する手法 **ChemTS** について詳しく紹介している。この手法では、**SMILES** と呼ばれる分子の文字列表現を生成するが、前半部分に関しては、モンテカルロ木探索で生成し、残りは、**RNN** で補完することで分子生成を行う。また、複数の分子生成を並列で行う **Virtual Loss** と呼ばれるモンテカルロ木探索の並列化手法についても紹介している。

第四章では、**NMR-TS** について詳しく紹介している。この手法の入力は、未知分子の **NMR** スペクトル（ターゲットスペクトル）である。**NMR-TS** は、4つのステップを繰り返すことで、分子同定を行う。第一に、**ChemTS** を用いて分子を一つ生成する。第二に量子化学計算を用いて生成された分子のスペクトルを計算する。第三に、**Wasserstein** 距離を用いて、ターゲットとの距離を評価する。最後に、計算された距離を **ChemTS** のパラメータにフィードバックすることで、ターゲットに近いスペクトルを持つ分子が生成されやすくする。量子化学計算に関しては、**Universal Force Field(UFF)**を用いて三次元配座を推定した上で、**B3LYP/3-21G\***レベルの密度汎関数理論（**DFT**）を用いている。また、**Wasserstein** 距離を用いることで、**NMR** スペクトルを離散化することを避けて、より詳細な誤差評価を可能にしている。さらに、過去に行ったモ

ンテカルロ木探索の履歴を、探索を開始する前に導入しておくことで、探索の高速化を図る **Trie Enhancement** という手法を提案し、実装している。**NMR-TS** の評価実験においては、**PubChemQC** データベースから、**C,H,N,O** のみで構成される分子で、かつ、電荷をもたない 9866 個の分子セットを作成し、**Recurrent Neural Network** の学習を行った。また、学習セットに含まれない 9 個の分子に関してターゲットスペクトルを計算した。各ターゲット分子に対して、20 コアを用いて 100 時間 **NMR-TS** を実行した結果、9 個のうち、6 個に関しては、正確に同定することができた。また、2 個に関しては、訓練セット中で最も近い分子（ベースライン）よりも、ターゲットに近い分子を発見することができた。

第五章は、**NMR** スペクトル同士の距離尺度に関する議論を行っている。**NMR-TS** では、**Wasserstein** 距離を用いたが、さらなる複雑な分子や、実験的に得られた雑音を含むスペクトルに対応するため、**Metric learning** による距離尺度の学習についても検討している。第六章では、これまでの内容をまとめた上で、将来の発展の方向性について示している。特に、環境サンプルなど、多数の分子がサンプルに含まれている場合への対応について議論している。

本論文の内容は、他の研究者との共同研究であるが、論文提出者が主体となって、手法提案及び検証を行ったもので、論文提出者の寄与が十分であると判断する。

よって本論文は博士（科学）の学位請求論文として合格と認められる。

以上 1889 字