

令和 6 年 6 月 26 日現在

機関番号：32829

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K03727

研究課題名（和文）Boolean多項式の連立方程式の公式による多重ゼータ値と計算機技術の融合研究

研究課題名（英文）Research on multiple zeta value and calculation technology by a formula of systems of Boolean polynomial equations

研究代表者

町出 智也（Tomoya, Machide）

東京国際工科専門職大学・情報工学科・講師

研究者番号：60614526

交付決定額（研究期間全体）：（直接経費） 1,800,000円

研究成果の概要（和文）：本研究は、数学に現れる多重ゼータ値と計算機科学における充足可能性問題（Satisfiability problem, SAT）の間に存在する共通の性質を見出し、融合研究を目指した。後者から前者の応用として、多重ゼータ値の関係式のランク計算に成功した。前者から後者の応用として、Boolean多項式の連立方程式の公式を通じたアルゴリズムの発展に貢献した。応用の際は自然言語処理の知識も利用した。

研究成果の学術的意義や社会的意義

本研究では数学の多重ゼータ値と計算機科学のSAT問題の結果を互いに応用した。また応用の際は自然言語処理の知識と経験を用いた。このことは研究の他分野連携を促進する。異なる分野の連携は思いもかけない発展につながる可能性が高く、昨今の大規模言語モデルの隆盛とともに、今後さらなる融合が考えられる。実際、多重ゼータ値の線形関係式の整数係数を mod 2 の条件下（つまり真偽の2値の条件下）で考察した場合、不可思議な法則があることが発見された。

研究成果の概要（英文）：In this research, we try to collaborate with subjects of mathematics and Computer science: the subject of mathematics is the multiple zeta value (MZV) in number theory, and that of Computer science is the Boolean satisfiability problem (SAT) in propositional logic. As an application from the latter to the former, we were able to calculate ranks of linear relations of MZVs. As an application from the former to the latter, we developed the algorithm via the formula for systems of Boolean polynomial equations. We also use the knowledge of natural language processing (NLP).

研究分野：解析的整数論、SAT、自然言語処理

キーワード：多重ゼータ値 Boolean多項式 ガウスの消去法 充足可能性問題(SAT) 彩色問題 Alon-Tarsi多項式

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

多重ゼータ値はリーマンゼータ関数の特殊値の一般化であり、1990 年以降、数学や物理学における様々な分野に現れ、盛んに研究されている。多重ゼータ値の線形関係式から構成される整数行列のランク計算は、多重ゼータ値のベクトル空間の次元と関連して重要だが、行列のサイズが指数的に増大するため、計算が非常に困難である。一方、人工知能および計算機工学における最も基本的な問題として、充足可能性問題 (satisfiability problem, SAT) がある。近年、大規模な SAT 問題を高速に解くアルゴリズムが開発され、様々な実世界への応用が拡大している。

多重ゼータ値の特徴として、非常にたくさんの線形関係式が存在する。更に関係式から定義されるベクトル空間には、特殊線形群上のモジュラー形式やフィボナッチ数的な数列と関連する、という事実がある。多重ゼータ値の線形関係式を全て生成すると予想されている関係式のクラスはいくつかあるが、その一つに“正規化ダブルシャッフル関係式”がある[1]。そのクラスは全ての関係式の生成という予想に加えて、「(a)整数係数、(b)引数の組に対して一つの線形関係式が対応、(c)個数が比較的少ない」という(特に計算機実験に有用な)性質を満たす。なお多重ゼータ値の関係式は重さ(引数の総和)で分類されるのが慣例である。それは、重さの異なるリーマンゼータ関数の特殊値は一時独立である、という経験則に基づく。

SAT は人工知能および計算機工学における最も基本的な問題として、システム検証、スケジューリング問題、定理証明など、様々な用途に応用されている。0 と 1 からなる 2 元体 $F_2 (= \{0, 1\})$ 上の足し算“+”は、論理演算“xor”と同値なため、「0 を False、1 を True」とみなすことにより、 F_2 上の線形関係式は命題論理式に変換できる。例えば、「 $x + y = 0$ 」は「 $(\neg x \vee y) \wedge (x \vee \neg y) = \text{True}$ 」に対応する。この変換を用いれば理論上では、整数係数を F_2 上に還元することにより、SAT 問題の解法に関する技術の利用が線形関係式に対して可能になる。しかし SAT では“or”と“and”の論理演算が想定されており、足し算に対応する“xor”は扱いにくい。そのため、SAT の技術を多重ゼータ値の線形関係式に応用するためには工夫が必要なことが想定される。例えば、“or”と“and”の論理演算における技術を“xor”ベースに翻訳すること、また申請者が発見した Boolean 多項式の連立方程式の公式[2]を使う、などである。

なお申請者は、以前所属していたプロジェクトにおいて、自然言語処理(natural language processing, NLP)の共同研究のため wikipedia の文書データに関する実験をしていた。機械学習や深層学習などを含む NLP の技術を多重ゼータ値に応用することは興味深い主題である。例えば、wikipedia のページのタイトルと単語の出現頻度を、正規化ダブルシャッフル関係式の上記の(b)の性質から、多重ゼータ値の引数の組と関係式に出現する多重ゼータ値の出現頻度とを対応させ、自然言語処理の技術を応用することが考えられる。しかし一般に、NLP の技術はデータにノイズがあることが想定されており、ノイズのないデータを扱う数論とは相性が悪い、という結論にその当時は至った。一方、SAT には矛盾ベースの探索技術 (Conflict Driven Clause Learning, CDCL) があった。SAT は、変数の値が「真」と「偽」の 2 値であり出力結果も「解あり」と「解なし」の 2 つのためノイズがなく、整数論と(特に 2 元体の条件下の元で)相性が良い。そのため本研究では SAT の技術を優先的に考察の対象とした。ただし NLP の実験の経験や技術も随所に活かすことも想定していた。

2. 研究の目的

本研究の目的は、正規化ダブルシャッフル関係式から得られる整数行列のランクの下限を、SAT や計算機工学の技術と申請者によって発見された Boolean 多項式の連立方程式に関する公式を利用し、既存の結果を上回りかつ新しい方法で計算することである。そして公式を通して、多重ゼータ値と SAT やその関連技術の間に存在する共通の性質を見出し、数学と計算機科学の融合研究を目指す。

多重ゼータ値の行列のランクの計算は、1990 年代から行われており、初期の頃はグレブナ基底の観点から線形構造だけではなく代数構造を考慮した実験もあった。金子-野呂-鶴巻[3]は、正規化ダブルシャッフル関係式から得られる行列のランクをガウスの消去法で調べ、ランクが予想の値以上になることを、重さ $k = 20$ まで確認した。行列サイズが 2^{k-2} で増加するため、「整数係数を mod $p (= 16381, 31991)$ する」「前処理により行列サイズを縮小する」「ガウスの消去法を並列化する」等の工夫をしている。なお、主張が等号ではなく以上となる理由は、mod p を取ることにより、ランクの下限しか示せなくなるためである。その計算では、重さ「18, 19, 20」の場合の計算で、それぞれ約「1 時間 20 分, 9 時間, 67 時間」の時間が掛かっていた。

本研究では上記の[3]の実験を基準に、以下の項目を目的とした：

1. [2]や[4]で提案した技術をアルゴリズムを線形連立方程式の場合に応用または拡張。
2. 応用または拡張したアルゴリズムを使用し、行列のランクを計算。

3. 計算結果を反省し、アルゴリズムの修正点と改良点を考察。そして、1 に戻る。

項目 1 において応用や拡張に使用するアルゴリズムは、[2]の二分木の構造を持つ公式を利用するアルゴリズムや、[4]で考察されたガウスの消去法と矛盾ベースの探索技術(CDCL)を組み合わせたアルゴリズムである。具体的な研究の方法は次の章で記述する。

3. 研究の方法

本研究では、[2]にある Boolean 多項式の公式や[4]のガウスの消去法と CDCL が組み合わされたアルゴリズムを軸にし、以下のような方針で行列のランクを計算した。

まず前述の通り、公式と SAT に代表される計算機工学の技術を利用するため、「整数係数を mod 2」して、つまり、体 F_2 の行列に帰着して考察した。また CDCL の他に変数選択の技術 [5] (Variable State Independent Decaying Sum, VSIDS) を参考にした。更に並列処理も組み込んだ。プログラム言語は Python の C 言語化に当たる Cython で書き換えた。実験マシンの用意は SAT の専門家 ([4]の共著者の藺部知大氏) に手伝ってもらった。変数選択の技術はアルゴリズムの効率化に強く影響するため、多重ゼータ値の出現頻度などの統計情報も考慮した。

重さ k が大きくなると関係式の個数も多くなり、データの保存の仕方に工夫を要する。そのため例えば、大量のデータを扱う NLP の手法に倣いデータベースソフトウェア mysql を使用する、バイナリファイル化のためのモジュール pickle を使用する、などの工夫をして対処した。また[2]の Boolean 多項式の公式は二分木の構造があるため、お姉さん問題に代表される決定二分木の方法 (ZDD) [6] の応用も当初は試みた。

4. 研究成果

2020 年度は、多重ゼータ値の整数行列のランク計算のアルゴリズムの改良を行った。具体的には、重さ 21 の場合のランクを計算した。また計算速度を向上させた。計算結果のランクの値は予想通りとなり、既に知られている、多重ゼータ値のベクトル空間の次元の上限と一致した。上限を与える具体的な生成元も知られているが、それは係数が有理数体の場合であった。今回の実験では、mod 2 の条件下でもそれが生成元となること(つまり 2 元体上 F_2 でも行列のランクが退化しないこと)がわかった。上記の結果以外では、Boolean 多項式の連立方程式の公式を使用したアルゴリズムを研究した。公式には二分木(BinaBinary Decision Diagram, BBD)の構造があるため、当初はその応用技術である ZDD(Zero-suppressed BDD)との関連を考察しようとしていた。しかし、グレブナ基底の基本的な技術である多変数のユークリッドの互除法がより有用でありそうなることがわかった。理由は互除法は足し算があっても影響はないが、ZDD では xor を使うことになり取り扱いが難しくなるからである。(2 元体 F_2 における足し算は論理演算の xor になることに注意する。)ZDD より互除法の方がより代数的であり申請者にとって取り組みやすいため、そちらを組み込んだアルゴリズムを部分的に作成した。

2021 年度は、多重ゼータ値の行列のランク計算のアルゴリズムを改良して、重さ 22 の場合のランクを計算した。更に深さで階層化した 2 元体 F_2 上のベクトル空間の基底も同時に計算した。階層化したベクトル空間ではより精密な計算を必要とし困難が予想されたため当初は計画していなかったが、ガウスの前進消去を発展させることにより成功を収めた。成功の理由は適切な列の順番が発見できたことに由来する。その基底は本質的に Hoffman 基底である。有理数係数の場合は Broadhurst-Kreimer 予想により Hoffman 基底とは無関係と考えられているが、2 元体の場合では異なる振る舞いをするのがわかった。(なお Broadhurst-Kreimer 予想は階層化された次元が保型形式の空間の次元と関連する、という予想である。)また、[2]の Boolean 多項式の連立方程式の公式が、グラフの彩色問題に現れる Alon-Tarsi 多項式と関連することがわかった。(Alon-Tarsi 多項式の存在は小関健太氏にご教示いただいた。)各頂点に約 $(\log k)$ 個の変数を対応させることにより k 彩色問題を $(\log k)$ 次の Boolean 多項式の連立方程式に変換できる(ただし \log の底は 2 とする)。そして Boolean 多項式の連立方程式はたった一つの多項式で表現できる。その多項式が Alon-Tarsi 多項式と関連する。二つの多項式には次の異なる箇所がある。Alon-Tarsi 多項式は整数係数であり彩色可能性が「多項式が 0 にならない」ということと同値である。一方 Boolean 多項式では、値が 1(真)と 0(偽)しかなく「0 にならない」は「1 になる」という意味と同じため、彩色可能性が「多項式が 1 になる」と同値になる。そのため、Alon-Tarsi 多項式の場合は探索的に解を探す必要があるが、Boolean 多項式の場合は純代数的に考察できる。上記の研究は両方とも当初の計画で想定していない事柄ではあるが、予想外の発見で今後の発展が期待できるため、当初の予定を変更して優先して研究した。

2022 年度以降は、前年度の多重ゼータ値の行列のランク計算の結果をまとめた。具体的には研究集会や小規模なセミナーで計算結果を発表して改善点を洗い出し、論文にまとめて投稿した。結果をまとめるにあたり、ダブルシャッフル関係式の二つの部分集合においても追加実験をした。一つは Minh-Jacob-Petitot-Oussous[7]により提案された部分集合で、有限ダブルシャッフル関係式と Hoffman 関係式の合併として定義される。もう一つは野呂-金子-鶴巻[3]による部

分集合で、上述の集合を制限して得られる。つまり前者は後者を集合として含む。両方とも有利係数上で実験し次元予想を担保することに成功していたが、2元体 F_2 上ではどのようになるか不明であった。(ここで以前の実験で、 F_2 上でもダブルシャッフル関係式が次元予想を担保することは既にわかっていること、そして予想の次元はフィボナッチ数列のような性質を満たすことに注意する。) 追加の実験の結果、両者とも次元予想を担保しないことがわかった。しかし Minh 氏等の関係式の集合のランク計算では、フィボナッチ数列に非常に近い性質があることがわかった。これは F_2 上のダブルシャッフル関係には有理数体上の場合と異なった(不可思議な)性質があることを示唆している。なお野呂氏等の関係式の集合のランクに法則性があるかどうかは発見できていない。本研究では自然言語処理の技術も参考にしているが、その関連として、共著の自然言語処理の論文が受理されたことを付記しておく。また、Boolean 多項式の連立方程式の公式を使ったアルゴリズムの研究と Alon-Taris 多項式との関連の研究も行った。

参考文献

- [1] K. Ihara, M. Kaneko, and D. Zagier, Derivation and double shuffle relations for multiple zeta values, *Compos. Math.*, 142 (2006), 307-338.
- [2] T. Machide, A formula for systems of Boolean polynomial equations and applications to computational complexity, arXiv:1907.09686, 2019.
- [3] M. Kaneko, M. Noro, K. Tsurumaki, On a conjecture for the dimension of the space of the multiple zeta values, *Software for algebraic geometry*, 47-58, IMA Vol. Math. Appl., 148, Springer, New York, 2008.
- [4] T. Machide and T. Sonobe, Determination method influenced by SAT solver for the full rankness of a matrix (in Japanese), *JSAI* 32, 2018.
- [5] H. Nabeshima and T. Soh, Principles of Modern SAT Solvers (in Japanese), *JSAI* 25, 2010.
- [6] S. Minato, Efficient Graph Enumeration and Indexing Using BDDs/ZDDs (in Japanese), *ORSJ* 57, 597-603, 2012.
- [7] H. N. Minh, G. Jacob, M. Petitot and N. E. Oussous, Aspects combinatoires des polylogarithmes et des sommes d'Euler-Zagier, *J. E´lectr. S´em. Lothar. Combin.* 43 (2000), Art. B43e, 29 pp.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 3件/うち国際共著 2件/うちオープンアクセス 3件）

1. 著者名 Bollegala Danushka, Otake Shuichi, Machide Tomoya, Kawarabayashi Ken-ichi	4. 巻 -
2. 論文標題 A Neighbourhood-Aware Differential Privacy Mechanism for Static Word Embeddings	5. 発行年 2023年
3. 雑誌名 In Findings of the Association for Computational Linguistics: IJCNLP-AAACL 2023 (Findings)	6. 最初と最後の頁 65-79
掲載論文のDOI（デジタルオブジェクト識別子） 10.18653/v1/2023.findings-ijcnlp.7	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Machide Tomoya	4. 巻 47
2. 論文標題 COMPUTATIONS ABOUT FORMAL MULTIPLE ZETA SPACES DEFINED BY BINARY EXTENDED DOUBLE SHUFFLE RELATIONS	5. 発行年 2023年
3. 雑誌名 Tsukuba Journal of Mathematics	6. 最初と最後の頁 83-111
掲載論文のDOI（デジタルオブジェクト識別子） 10.21099/tkbjm/20234701083	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 町出智也	4. 巻 2238
2. 論文標題 ダブルシャッフル関係式から導かれるバイナリ行列	5. 発行年 2022年
3. 雑誌名 数理解析研究所講究録 2238「多重ゼータ値の諸相」	6. 最初と最後の頁 47-57
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Danushka Bollegala, Tomoya Machide, Ken-ichi Kawarabayashi	4. 巻 -
2. 論文標題 Query Obfuscation by Semantic Decomposition	5. 発行年 2022年
3. 雑誌名 Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)	6. 最初と最後の頁 6200-6211
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計8件（うち招待講演 1件 / うち国際学会 3件）

1. 発表者名 Danushka Bollegala
2. 発表標題 A Neighbourhood-Aware Differential Privacy Mechanism for Static Word Embeddings
3. 学会等名 In Findings of the Association for Computational Linguistics: IJCNLP-AAACL 2023 (Findings) (国際学会)
4. 発表年 2023年

1. 発表者名 町出智也
2. 発表標題 Boolean多項式の連立方程式の彩色問題への応用
3. 学会等名 日本数学会2023年度年会
4. 発表年 2023年

1. 発表者名 町出智也
2. 発表標題 矛盾探索を土台にしたガウスの前進消去法について
3. 学会等名 日本応用数理学会2022年度年会
4. 発表年 2022年

1. 発表者名 Danushka Bollegala
2. 発表標題 Query Obfuscation by Semantic Decomposition
3. 学会等名 Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022) (国際学会)
4. 発表年 2022年

1. 発表者名 町出智也
2. 発表標題 バイナリ拡張ダブルシャッフル関係式に関する形式的多重ゼータ空間の次元の計算
3. 学会等名 2022年度RIMS共同研究(公開型)多重ゼータ値の諸相(招待講演)(国際学会)
4. 発表年 2022年

1. 発表者名 町出智也
2. 発表標題 Boolean 多項式の連立方程式の彩色問題への応用
3. 学会等名 日本数学会2022年度年会
4. 発表年 2022年

1. 発表者名 町出智也
2. 発表標題 Boolean 多項式の連立方程式によるリスト彩色問題の解法
3. 学会等名 2021年度応用数学合同研究集会
4. 発表年 2021年

1. 発表者名 町出智也
2. 発表標題 ブール多項式の連立方程式に関する公式について(On a formula for systems of Boolean polynomial equations)
3. 学会等名 第20回広島仙台整数論集会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------