

令和 5 年 6 月 6 日現在

機関番号：32675

研究種目：若手研究

研究期間：2019～2022

課題番号：19K20350

研究課題名(和文) 局所的重要性に着目した知識発見基盤技術の開発

研究課題名(英文) Fundamental techniques for knowledge discovery based on local importance

研究代表者

和佐 州洋 (Wasa, Kunihiro)

法政大学・理工学部・講師

研究者番号：00781337

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：期間全体を通じてアルゴリズム構築のための基盤的な手法について、その端緒を得ることができた。一方で、「大事さ」に関する問題の収集は、想定目標には到達できなかった。その具体的な原因は、コロナ禍による人的交流の極端な減少による実問題の収集のハードルの高さが考えられる。昨今、機械学習における説明可能性のみならず、非明示的な指標を背後に持った k-best 列挙などのアルゴリズムに対して、一定程度、その説得力を持つことが要請されている。本課題終了後も、継続してこのような問題について取り組んでいく。

研究成果の学術的意義や社会的意義

列挙問題、つまり、一定条件を満たした解を全て列挙する問題はさまざまな分野において見られる基礎的な問題である。本研究成果の学術的意義は、本課題の一つの課題である種々の問題に対する列挙アルゴリズムの構築を実際に行った点である。一方で、このアルゴリズムを利用して「解きたい問題」においては何に着目したら良いかを導き出す、というもう一つの課題については成果を上げられなかった。この点については、今後同様の問題意識を持って取り組み成果を上げていきたい。

研究成果の概要(英文)：Throughout the entire period, we were able to get a basic idea for a framework for developing enumeration or reconfiguration algorithms. However, the other goal, that is, revealing "what is the importance measure of problems" did not achieve. The specific reason for this may come from the difficulty of meeting people in the other field due to the extreme decrease in human interaction caused by the Corona disaster. Recently, when we develop an algorithm, users ask us to clarify "why such a solution is outputted?" This kind of explainability is highly demanded such as k-best enumeration, which have non-explicit indicators behind them as well as machine learning. We will continue to address this topic after the completion of this project.

研究分野：離散アルゴリズム

キーワード：列挙アルゴリズム データマイニング 近傍領域

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

大量に蓄積されたデータにより、私たちは新たな知識を発見できるようになった。一方で、その大量さは同時にこれまで利用できたデータマイニングアルゴリズムを適用できない規模となっている。この状況のもとで、我々は何もできないのだろうか？本研究は、そのような状況でも新しいアプローチによってこれまで解くことのできなかった規模の問題を解いていくモチベーションのもとで研究をスタートした。

2. 研究の目的

本研究は、「近傍領域」と呼ばれる局所的な構造に着目することで、この状況を打破しようというものである。現実の問題を解く上で重要なことは、ノイズなどを想定しない理想的な状況のもとでのモデル構築ではなく、シンプルな視点で問題を観察し、その観察から得られた情報、いわば中間生成物を利用して、元の問題を解くことである。この視点に基づき、本研究では、着目したデータの重要性を高くする近傍領域を抽出する技術の開発を目標とする。

3. 研究の方法

大量に蓄積されたデータにより、私たちは新たな知識を発見できるようになった。一方で、その大量さは同時にこれまで利用できたデータマイニングアルゴリズムを適用できない規模となっている。本研究では、「近傍領域」と呼ばれる局所的な構造に着目することで、この状況を打破しようというものである。具体的には、次の二つの課題を設定した。これらの課題を克服することで、知識発見技術のさらなる発展を狙う。

課題1: 現実の問題に対する適切な「大事さ」に関するパラメータのモデル化

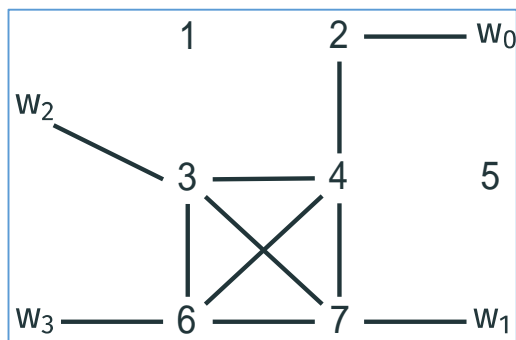
課題2: 課題1で得られたパラメータを元に、近傍領域を効率よく計算するアルゴリズムの開発

4. 研究成果

課題1については、顕著な成果を上げることができなかった。一方で、課題2について、予備的な結果ではあるが幾つかの列挙アルゴリズム、組合せ遷移問題に対するアルゴリズム、及び困難性を証明することに成功した。具体的には、マッチング、グラフ彩色、誘導木、葉の数や度数に制約のある全域木、 k -縮退グラフ、密グラフ、支配集合、阿弥陀籤などのグラフをはじめとする離散構造を対象にした成果を得ることができた。以下では、それらの中でも3つの成果について述べる。

A. シュタイナーグラフの列挙

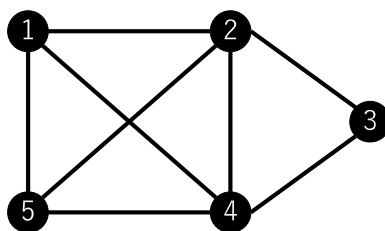
グラフとはさまざまな現実データを表現できる離散構造の一つである。下の図([Conte, et al. MFCS2019]より引用)は、グラフの例である。グラフは点(数字とアルファベット)と線からなっており、身近な例としては鉄道網などがこのような形式で表現される。ここでの、グラフの点を2種類に分ける。一つはターミナル頂点(w_0, \dots, w_3 で表現)、もう一つはシュタイナー頂点(数字で表現)である。ここでシュタイナーグラフとは、ターミナル頂点を全て含み、かつ、全てのターミナル頂点間に経路が存在するようなグラフをいう。下図の薄字になった部分を除いたものがシュタイナーグラフとなっており、このうち、どのシュタイナー頂点を除いてもシュタイナー性は満たされない(このような性質を極小性を満たすとい



う)。本研究の主結果の一つは、このようなシュタイナーグラフでも極小性を満たしたものを高速に発見し、列挙するアルゴリズムを構築した点である [Conte, et al. MFCS 2019] [Kobayashi, et al., PODS 2022]。さらに、この問題のある種の実例が、長年の未解決問題と同等以上に難しいことも証明した [Conte, et al. MFCS 2019]。これらの成果は、理論重視の国際会議及びデータベースの応用重視の国際会議で採択されており、学術的な成果は大きい。また、「ターミナル頂点は重要である」という観点に立つならば、本研究の課題 1 において、このような頂点が重要視されるケースでは有用なアルゴリズムである。

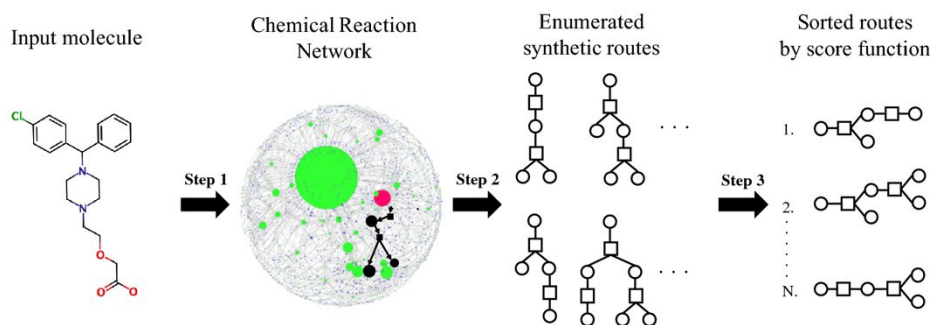
B. オイラー路の列挙

オイラー路とは、グラフの辺をちょうど 1 回ずつ辿って得られる路のことを言う。下の例では、1 2 3 4 1 5 2 4 5 と辿ると、全ての辺をちょうど 1 度ずつ通るような経路になっている。このオイラー路はグラフ理論においても重要な研究対象であり、その存在性の必要十分条件はオイラーによって与えられた。一方で、列挙については、アルゴリズムがほとんど知られてこなかった。具体的には、菊池による一つあたり線形時間遅延の列挙しか与えられていなかった。本研究では、これらオイラー路を一つあたり定数時間で列挙するアルゴリズムを与えた。このアルゴリズムは、全体として解の総数に対して線形時間で動作していることから理論的に最適なアルゴリズムとなっている。本成果は Theoretical Computer Science に受理されており、理論計算機科学分野での重要な貢献となっている。



C. 化合物合成経路の列挙

化学反応ネットワークにおける最適ルートと代替ルートを検討するためには、多くの化学合成ルートを得る必要がある。しかし、これまでのプログラムの多くは、高々数本の最適なルートのみを出力するものであった。本論文では、いくつかの条件を満たした反応経路を全て出力するプログラム “CompRet” を提案し、予備実験において、既知の抗ヒスタミン薬であるセチリジンの代替ルートを発見することに成功した。本研究では主に下の図 ([Shibuya et al., J. Cheminformatics, 2020] より引用) の “Enumerated synthetic routes” の部分でアルゴリズム開発に貢献した。本成果は、Journal of Cheminformatics に採択されている。



5. 主な発表論文等

〔雑誌論文〕 計13件（うち査読付論文 13件 / うち国際共著 4件 / うちオープンアクセス 2件）

1. 著者名 Kurita Kazuhiro, Wasa Kunihiro	4. 巻 923
2. 論文標題 Constant amortized time enumeration of Eulerian trails	5. 発行年 2022年
3. 雑誌名 Theoretical Computer Science	6. 最初と最後の頁 1~12
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.tcs.2022.04.048	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 -
1. 著者名 Kobayashi Yasuaki, Kurita Kazuhiro, Wasa Kunihiro	4. 巻 13453
2. 論文標題 Polynomial-Delay and Polynomial-Space Enumeration of Large Maximal Matchings	5. 発行年 2022年
3. 雑誌名 Proc. of The 48th International Workshop on Graph-Theoretic Concepts in Computer Science 2022	6. 最初と最後の頁 342~355
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-15914-5_25	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 Kobayashi Yasuaki, Kurita Kazuhiro, Wasa Kunihiro	4. 巻 -
2. 論文標題 Linear-Delay Enumeration for Minimal Steiner Problems	5. 発行年 2022年
3. 雑誌名 Proc. of The 41st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems	6. 最初と最後の頁 301~313
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3517804.3524148	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 Kurita Kazuhiro, Wasa Kunihiro, Uno Takeaki, Arimura Hiroki	4. 巻 874
2. 論文標題 A constant amortized time enumeration algorithm for independent sets in graphs with bounded clique number	5. 発行年 2021年
3. 雑誌名 Theoretical Computer Science	6. 最初と最後の頁 32~41
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.tcs.2021.05.008	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 Kurita Kazuhiro, Wasa Kunihiro, Arimura Hiroki, Uno Takeaki	4. 巻 303
2. 論文標題 Efficient enumeration of dominating sets for sparse graphs	5. 発行年 2021年
3. 雑誌名 Discrete Applied Mathematics	6. 最初と最後の頁 283 ~ 295
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.dam.2021.06.004	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 Ito Takehiro, Iwamasa Yuni, Kobayashi Yasuaki, Nakahata Yu, Otachi Yota, Wasa Kunihiro	4. 巻 13025
2. 論文標題 Reconfiguring Directed Trees in a Digraph	5. 発行年 2021年
3. 雑誌名 Proc. of The 27th International Computing and Combinatorics Conference (COCOON 2021)	6. 最初と最後の頁 343 ~ 354
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-89543-3_29	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 Eto Hiroshi, Ito Takehiro, Kobayashi Yasuaki, Otachi Yota, Wasa Kunihiro	4. 巻 13174
2. 論文標題 Reconfiguration of Regular Induced Subgraphs	5. 発行年 2022年
3. 雑誌名 Proc. of The 16th International Conference and Workshops (WALCOM 2022)	6. 最初と最後の頁 35 ~ 46
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-96731-4_4	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 Nicolas Bousquet, Takehiro Ito, Yusuke Kobayashi, Haruka Mizuta, Paul Ouvrard, Akira Suzuki, Kunihiro Wasa	4. 巻 219
2. 論文標題 Reconfiguration of Spanning Trees with Degree Constraint or Diameter Constraint	5. 発行年 2022年
3. 雑誌名 Proc. of The 39th International Symposium on Theoretical Aspects of Computer Science (STACS 2022)	6. 最初と最後の頁 15:1 ~ 15:21
掲載論文のDOI (デジタルオブジェクト識別子) 10.4230/LIPIcs.STACS.2022.15	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 該当する

1. 著者名 Katsuhisa Yamanaka, Takashi Horiyama, Kunihiro Wasa	4. 巻 859
2. 論文標題 Optimal reconfiguration of optimal ladder lotteries	5. 発行年 2021年
3. 雑誌名 Theoretical Computer Science	6. 最初と最後の頁 57 ~ 69
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.tcs.2021.01.009	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 Ryosuke Shibukawa, Shoichi Ishida, Kazuki Yoshioze, Kunihiro Wasa, Kiyosei Takasu, Yasushi Okuno, Kei Terayama, Koji Tsuda	4. 巻 12
2. 論文標題 CompRet: a comprehensive recommendation framework for chemical synthesis planning with algorithmic enumeration	5. 発行年 2020年
3. 雑誌名 Journal of Cheminformatics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s13321-020-00452-5	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 Marthe Bonamy, Marc Heinrich, Takehiro Ito, Yusuke Kobayashi, Haruka Mizuta, Moritz Muhlenthaler, Akira Suzuki, Kunihiro Wasa	4. 巻 838
2. 論文標題 Diameter of colorings under Kempe changes	5. 発行年 2020年
3. 雑誌名 Theoretical Computer Science	6. 最初と最後の頁 45 ~ 57
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.tcs.2020.05.033	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 該当する

1. 著者名 Alessio Conte, Mamadou Moustapha Kante;, Takeaki Uno, Kunihiro Wasa	4. 巻 303
2. 論文標題 Maximal strongly connected cliques in directed graphs: Algorithms and bounds	5. 発行年 2021年
3. 雑誌名 Discrete Applied Mathematics	6. 最初と最後の頁 237 ~ 252
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.dam.2020.05.027	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 該当する

1. 著者名 Alessio Conte, Roberto Grossi, Mamadou Moustapha Kante, Andrea Marino, Takeaki Uno, and Kunihiro Wasa	4. 巻 118
2. 論文標題 Listing Induced Steiner Subgraphs as a Compact Way to Discover Steiner Trees in Graphs	5. 発行年 2019年
3. 雑誌名 Proc. of The 44th International Symposium on Mathematical Foundations of Computer Science (MFCS 2019)	6. 最初と最後の頁 73:1 ~ 73:14
掲載論文のDOI (デジタルオブジェクト識別子) 10.4230/LIPIcs.MFCS.2019.73	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1件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 小林 靖明, 栗田 和宏, 和佐 州洋
2. 発表標題 Polynomial Delay Enumeration for Steiner Problems
3. 学会等名 電子情報通信学会 コンピューテーション研究会
4. 発表年 2020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
---------	---------