

部分構造に基づく構造類似性を用いた特徴抽出システムとその応用

Feature Extraction System Using Similarity Measure based on Substructure Distribution Analysis

和田 貴久 ♡
稲積 宏誠 ♡

大野 博之 ♡

Takahisa WADA
Hiroshige INAZUMI

Hiroyuki OONO

データの多様化に伴い、複雑な構造を持つグラフデータを扱うための技術開発の必要性が高まっている。本稿では、対象とするグラフ集合に内在する特徴的な部分構造を抽出し、それを用いることによって定義できる構造類似性に注目し、類似度計算過程において背景知識の導入を可能とする拡張や、得られたクラスタリング結果の評価とその結果の有効な活用方法について検討する。さらに、これらの検討結果に基づき、対象グラフ集合からの特徴抽出システムを提案する。また、特にグラフ構造データの典型である化学構造データを分析対象とした適用例を示し、その有効性を示す。

The graph data with the complex structure increases more and more along with the diversification of the accumulated data. We proposed structure similarity and graph clustering in accordance with the feature of target graph sets, and have been examining the characteristic. In this paper, to introduce background knowledge in target field into similarity calculation, we extend this clustering algorithm and validate the evidence. And we propose new structural analysis system for target graph sets. This system is applied to create new graph structures for new molecular entity.

1. はじめに

近年、Web 上にはテキストデータやトランザクションデータ、グラフデータなど、さまざまな形式のデータが蓄積され、それらを活用しようと、多くのマイニング手法が研究されている。ただし、従来の解析対象の多くはテキストデータやトランザクションデータなどのデータそのものであり、構造情報を含むグラフデータに対するマイニング手法の研究は比較的新しい分野とい

える。しかしながら、近年このようなグラフデータを扱うマイニング手法については、精力的に研究が進められ、多くの有用なアルゴリズムが提案されている [2]。特に、対象とするグラフ集合に共通する部分構造抽出のためのアルゴリズムを用いて、知識発見のための多くの取り組みがなされている。我々も、特に Graph-Based Induction (GBI) 法 [?] や GBI 法の改良手法である Chunkingless Graph-Based Induction (Cl-GBI) 法 [5] に注目して、化学物質の特性を分析するためのツールの開発や、抽出された部分構造情報を用いた分類問題やクラスタリングへの応用について検討を行ってきた [1, 8]..

このような構造的な特徴の活用方法においては、与えられたグラフ集合のグラフごとに、グラフ中に含まれる連結部分構造を列挙し、それを数値化したものを利用するという考え方が有効である [6]。これを受けて、グラフ構造情報のより有効な分析を実現するために、グラフにおける構造上の類似性に注目し、分析対象とするグラフ集合全体の持つ構造上の特徴に基づくグラフ間の構造類似性が検討されてきた。例えば、グラフ構造の構造的な特徴付けに対しては、化学物質の構造的類似性の観点から TFS (Topological Fragment Spectra) によるデータマイニング手法が提案されている [7]。また、我々は対象とするグラフ集合から、ある条件に基づき抽出された特徴を利用する手法として、連結部分構造を包含関係に基づく階層化による木表現を行い、その階層関係に基づく構造的類似性を導入した分析手法 (Hierarchical Substructure Analysis: HSA) を提案した [1]。さらに、Cl-GBI 法によって抽出された部分構造が各グラフ中にどのように分布しているかを評価することで、対象とするグラフ集合中のグラフ間の構造類似性を評価する手法 (Substructure Distribution Analysis: SDA) を提案した [8]。この手法は、抽出された部分構造が各グラフ中にどのように分布しているかを定義し、その分布を比較することによって類似度を評価するものである。

本稿では、まず、SDA 法に基づく類似度計算の概要を示したのち、対象グラフ集合のもつ背景知識や評価に利用したい属性項目を類似度の導入を可能とする拡張、さらにそれに基づくクラスタリング手法である SDA クラスタリング法について述べる。次に、SDA 法に基づく対象グラフ集合の特徴抽出システムを提案する。このシステムは、類似度計算処理の自動化と複数のクラスタリングアルゴリズムが実装されており、クラスタリング結果の各クラスタの特徴評価や結果の活用を可能とする。また、システム要件としては、一般グラフに対して適用可能であるが、特にグラフ構造データの典型である化学構造データを取り上げ、SDA 法で求められる類似度の特性評価と部分構造抽出における最適パラメータについての検討を行う。さらに、新規化合物の合成を行う上での有用な知識の提示を含めた応用について展望する。

2. 部分構造情報に基づく構造類似性

2.1 SDA(Substructure Distribution Analysis) 法による類似度計算 [8]

SDA 法は、対象とするグラフ集合全体の性質を背景とし、そのなかでの相対的な類似性に基づく分析を実現することを目的とするものである。そのためには、対象とするグラフ集合全体を特徴づけることが必要である。SDA 法では、この特徴づけをグラフ集合全体のなかにある一定基準以上存在する部分構造の組合せ

♡ 学生会員 青山学院大学大学院理工学研究科博士前期課程 t-wada@ina-lab.it.aoyama.ac.jp

♡ 非会員 青山学院大学理工学部情報テクノロジー学科 oono@ina-lab.it.aoyama.ac.jp

♡ 正会員 青山学院大学理工学部情報テクノロジー学科 hiro@ina-lab.it.aoyama.ac.jp

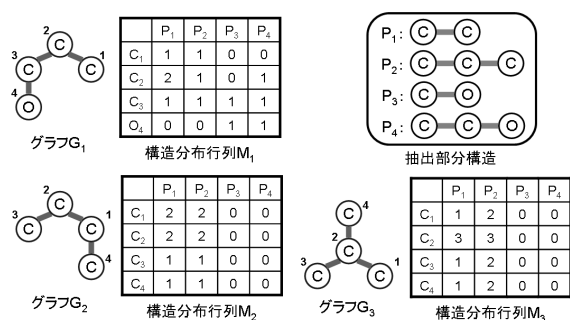


図1 構造分布行列の例

Fig. 1 Examples of substructure distribution matrices

を評価することにより実現できると考えた。すなわち、各グラフは、どのような部分構造をどのように保持しているかにより特徴づけられるという考え方に基づくものである。グラフを構成する一つ一つのノードが、どのような部分構造にどの程度含まれているかを情報として保持し、それをグラフの特徴量として定義し、グラフ間の構造的な類似性を評価することにより実現される。

そのためには、抽出漏れがなく、対象とするグラフ集合に共通する部分構造を効率的に抽出する必要がある。そこで、部分構造抽出アルゴリズムとしては、CI-GBI法を採用することとした。CI-GBI法の特徴は、頻度あるいはその他の基準を満足する部分構造を、重複を許すことで漏れなく抽出することが可能であることに加えて、一定条件を満足するもののなかで、戦略的に優先度をつけて抽出することも可能であるからである[?]。CI-GBI法により抽出された部分構造は、CI-GBI法の特性から、チャンク過程を示す履歴情報を用いて、それぞれの包含関係を示す情報が保持されている。また、分析対象であるグラフの各ノードには、擬似ノード生成過程から、抽出されたどの部分構造の構成要素となっているかという情報が保持されている。これらが、構造類似性を評価するうえでの基本情報となる。この部分構造を用いることによって、各グラフは、それを構成するノードと部分構造との関係に基づいて表現することができる。この関係を実現させた行列を構造分布行列(以後、SD行列と略す)と呼び、グラフの構造上の特徴が表現されているとみなす。図1にSD行列の例を示す。これは、対象とするグラフ集合全体を化学物質としたときに、抽出された部分構造が $p_1: C-C$, $p_2: C-C-C$, $p_3: C-O$, $p_4: C-C-O$ であったときの、グラフ G_1 , G_2 , G_3 の特徴表現である。

SDA法に基づく類似度は、このSD行列の差分を用いて定義する。まず、任意の2つのグラフ中に含まれる共通ラベルを持つノード集合ごとに、各ノード間の類似性を定義し、次に、そのノード間の類似性を用いてグラフ間の構造類似性を定義する。このとき、ノード間の類似性を計算するノードの組み合わせは、各ノードの部分構造との関連数の分布、すなわち、SD行列の横方向のベクトルの要素の分布が似ているもの同士を組み合わせる。この組み合わせごとに各要素の一致している数と一致していない数との差を求め、それぞれに対して各部分構造の重みを掛け合わせる。このときの重みは、部分構造の大きさとなる。この一致数と不一致数の割合をノード間の類似度とし、ノード間類似度の累計をグラフ間類似度とする。図1を例にすると、 $G_1 \cdot G_2$ の類

似度は0.607, $G_1 \cdot G_3$ の類似度は0.441, $G_2 \cdot G_3$ の類似度は0.682となる。

2.2 グラフデータの特性情報の導入

SDA法の類似度計算では、元来、大きな部分構造を共通に含んでいる場合により類似性が高くなるように、重みとして部分構造のサイズを与えていた[8]。しかしながら、対象とする問題領域においては必ずしもサイズが大きければよいという考え方が成り立つとは限らない。例えば、化学の分野においては、特別な性質を持つ構造が存在し、グラフの性質決定に大きな影響を与える場合がある。そのような場合には、特別な性質を持つ構造を含む部分構造の重みを大きくするのが望ましいと考えられる。

そこで、詳細な説明は省略するが、後述するシステムにおいては、部分構造の重みを任意の値に変更可能にし、一定ルールに基づいた重み付けを行うという考え方を導入する。例えば、ある特性値 r が対象とする各グラフに与えられて、この r の値の近いグラフは、性質が似ているものとして関係付けたいとする。例えば、化学物質における活性値などはその典型例である。その場合には、抽出された部分構造ごとに、それを含むグラフの r の値の分布を計算し、偏りの大きさに応じて、当該部分構造に重み付けを行う。これは、特性を決定する性質の強いものをより重視することであり、類似度評価に導入するという目的に合致することになる。このように、部分構造の大きさだけでなく、グラフの持つ非構造上の特性を類似度計算に導入することを可能とした。

3. SDA法の応用

3.1 クラスタリング

クラスタリングアルゴリズムは、SDA法に基づく類似度を用いて、最短距離法、最長距離法、群平均法による階層的クラスタリングと k -medoid法による非階層的クラスタリングなど、目的に応じて使い分けることを想定する。また、クラスタサイズについては、対象データの背景知識や対象データの特徴分布、特に、すべての事例間の類似度行列から仮想的な二次元配置を主座標分析により求め、その結果を活用するなどが考えられるが、詳細な検討は今後の課題とする。

3.2 クラスタリング結果からの知識獲得

構造上の特徴を基にしたクラスタリングでは、構成されるクラスタの特徴を把握することは困難である。したがって、クラスタリング結果を有効に活用するためには、各クラスタの特徴をどのように表現するかという点が重要となる。そこで、得られたクラスタの特徴を説明するために、まず、クラスタ内のすべてのグラフに含まれる最も大きな部分構造と、その部分構造に接続する部分構造を抽出する。これらの部分構造は、それぞれ最大共通構造と差分構造とみなすことができる。この最大共通構造に注目すると、クラスタの中に含まれているグラフの構造の傾向を知ることができ、差分構造に注目することで、クラスタ内のグラフの違いだけに焦点を当てることができる。

3.2.1 最大共通構造の抽出

最大共通構造を抽出するためには、対象グラフ集合内のすべてのグラフに含まれる部分構造を改めて抽出する必要がある。これについては、CI-GBI法のパラメータとチャンク候補の選択基準を変更し、適用することで同一システムとして実現することができる。すなわち、含有率を100%とし、チャンク過程において

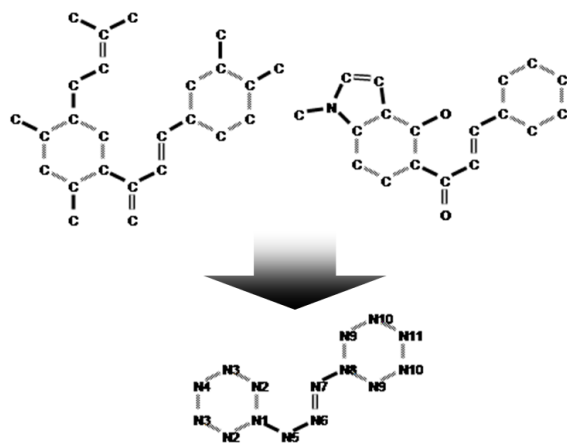


図2 最大共通構造の抽出と再ラベル付け

Fig. 2 Extraction and relabeling of the maximum common graph in each cluster

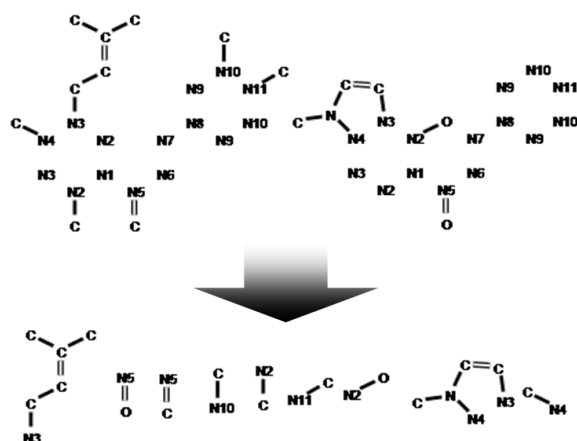


図3 抽出した差分構造の例

Fig. 3 Examples of difference-subgraphs

は、チャンク候補としてサイズの大きなものを優先することにより実現可能である。その結果、最後に得られた部分構造の中で、最も大きな部分構造を最大共通構造とする。また、最大共通構造内の各ノードに対して、位置関係を考慮した一意の再ラベル付けを行い、複数のグラフ内でのこの構造の配置を特定する。図2に2種類のグラフから最大共通構造を抽出し、それを再ラベル付けした例を示す。ここで、再ラベル化された $N_i (i = 1, 2, \dots, 11)$ は、元のグラフ中のラベルとユニークに対応付けられ、その情報が保持される。

3.2.2 差分構造の抽出

差分構造を抽出するためには、最大共通構造に接続する部分構造を抽出しなくてはならない。そこで、元のグラフに再ラベル化された情報を反映させ、グラフ内の最大共通構造に対応する部分のリンクをすべて除去する。ノードと残ったリンクのみのグラフからすべての連結部分構造を抽出することによって、差分構造を

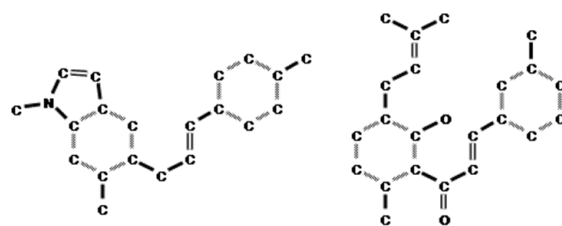


図4 新規グラフの例

Fig. 4 Examples of new composed-graphs

特定する。ただし、包含関係にある部分構造はより大きな部分構造を残し、小さな部分構造は除去する。同様に、対象となるグラフ集合内のすべてのグラフに対して部分構造抽出する。すべてのグラフから部分構造が抽出されたら、同型の部分構造が存在していないかを調べ、存在している場合は一方を除去する。

3.3 新規グラフ構造の自動生成

クラスタリング結果から各クラスタ内の最大共通構造と差分構造の情報を得ることができることを示した。その活用例としては、新規グラフ構造の生成が考えられる。これは、薬品などの新規化合物合成のためのシミュレーションに相当するものである。例えば、最大共通構造に対して、差分構造を再配置していくことで、複数のグラフの特徴を併せ持つような新たなグラフを生成することができる。既に、最大共通構造内のノードに関しては一意に識別できるように再ラベル付けがなされており、差分構造に関しても最大共通構造内のどのノードに接続すべきかが特定されているので、容易に配置することが可能である。ただし、一つのノードに対しては、一つの差分構造のみの配置を許すことにし、複数の差分構造が競合した場合、それぞれの別のグラフとして生成する。また、接続する差分構造の選択においては、差分構造がもともと含まれていたグラフの性質情報を参照することで、目的に合わせて任意に選択することも可能とすることができる。

4. SDA 法に基づく特徴抽出システム

部分構造の分布情報による分析方法である SDA 法を化学物質及びその他のグラフ構造データに適用可能な分析ツールを開発した。特に、SDA 法による分析の自動化と結果の可視化に基づき実装化されている。また、対象グラフ集合からの部分構造抽出を CI-GBI 法によるものとし、構造分布行列を用いた類似度計算を実現している。ただし、CI-GBI 法の各パラメータの設定、HSA 法による類似度計算、および、グラフの特性値を加味した重みに変更を可能とした。

クラスタリングアルゴリズムは、階層的クラスタリングである最短距離法、最長距離法、群平均法と、非階層的クラスタリングである k-medoid 法が実装されている。手法を選択し、分割するクラスタ数をしていてクラスタリングを行うことができる。クラスタリング結果の表示例を図5に示す。図5:Aと図5:Bでは、クラスタ内のすべてのグラフに含まれる共通構造とその共通構造に接続している差分構造を提示することが可能となっている。

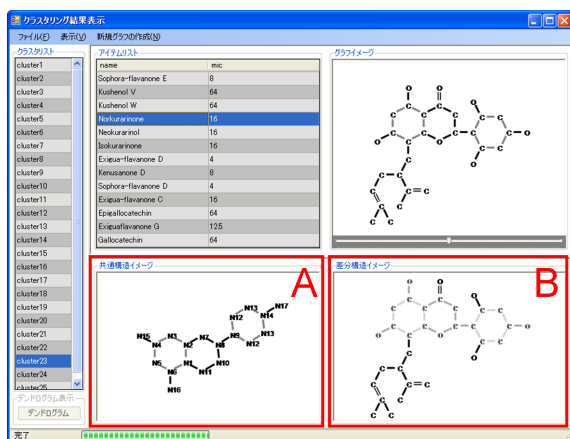


図5 クラスタリング結果の表示画面

Fig. 5 Display screen of Clustering Results

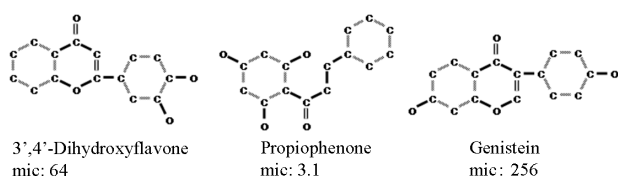


図6 フラボノイドの一例

Fig. 6 Examples of flavonoids data

5. 実験

5.1 概要

本実験では、まず、SDA 法により求められる類似度の特性の評価を行う。類似度計算に用いる部分構造集合を Cl-GBI 法のパラメータを変化させることで、複数パターン用意し、使用する部分構造集合が異なるときの類似度の違いを検証する。次に、グラフの特性情報を導入したことによる類似性の変化を検証する。最後に、特徴抽出システムを利用し、新規化合物の候補となるようなグラフ構造を生成する。

5.2 実験データ

実験データとして、図 6 に示すような抗菌活性物質のフラボノイド類 274 種類の化学構造データを用いる。これらのフラボノイド類の構造的特徴は、2つのベンゼン環が複数個の連結した炭素原子によって接続された構造を持つことである。ただし、複数の炭素原子の結合の種類は多様な組み合わせが存在する。この規則を満足した基盤構造に対して様々な置換基を配置することで生成されている。また、これらの化学構造データは、構造情報の他に、各化学物質の抗菌活性度の数値が属性として付加されている。この数値は、小さいほど高活性の物質であることを示す。

5.3 抽出部分構造の類似性への影響

SDA 法によるグラフ間の類似度の計算には、Cl-GBI 法によって抽出される部分構造を利用する。Cl-GBI 法は、チャンク処理の繰り返し数 (Level) と一度のチャンク処理で同時にチャンクを行う個数 (Beam) と部分構造のグラフに含まれる割合 (サポート) のパラメータによって、抽出される部分構造に違いが生じる。

表 1 Cl-GBI 法のパラメータと実行結果

Table 1 The numbers of substructures extracted from Graph set using Cl-GBI with various parameters

	サポート (%)	Level	Beam	抽出数	抽出時間 (秒)
A	10	3	3	38	6
B	10	5	5	131	32
	10	5	15	385	148
	10	10	10	483	650
	10	15	5	370	497
C	10	15	15	945	4373
D	20	5	5	91	44
E	20	5	15	314	174
	20	10	10	347	565
	20	15	5	255	734
	30	5	5	103	35
	30	5	15	203	142
F	30	10	10	283	736
	30	15	5	228	485

Level や Beam の値を大きくすることで、粒度の細かい抽出が可能となり、非常に密な構造分布行列が定義され、対象グラフ集合のグラフの特徴を高い精度で表現しているとみなすことができる。また、サポートを下げることで、大きな部分構造の抽出が可能とする。そのため、大きな部分構造を共通に持つものに、より高い類似性評価を行うことになる。これらを実現するためには、Level や Beam の値を大きくする必要があり、ある時点から急激に処理時間が長くなることが報告されている。

そこで、実際の応用を検討する際には、比較的粒度が粗く、低いサポートでより適正な類似度評価が実現されることが期待される。まず、サポートと Level と Beam のパラメータを変化させ、いくつかの部分構造群を抽出し、表 1 に示す。以下、サポート、Level、Beam の三つ組を C: (10, 15, 15) と表す。C は、最も粒度が細かく、本実験の類似度計算における基準値とみなすこととする。また、最も粒度が荒いのは、抽出数も最も少ない A: (10, 3, 3) である。A, C に加えて、B: (10, 5, 5) における類似度計算例を図 7 に示す。

この図で比較されているグラフ、3',4'-Dihydroxyflavone と Desmethylnobiletin は、2つのベンゼン環とその間にあるノードの配置が共通しているが、Isowighteone は、この部分のベンゼン環の配置が大きく異なっている。類似度を比較すると、B と C においては、3',4'-Dihydroxyflavone と Desmethylnobiletin が最も高く、A の場合では、3',4'-Dihydroxyflavone と Isowighteone の類似度が高くなっている。これは、大きな構造上の違いを発見できるか否かという点で、A の部分構造が、粒度の点での限界を示唆しているといえる。

さらに、D: (10, 5, 5), E: (10, 5, 5), F: (10, 5, 5) による類似度計算例を図 8 に示す。いずれの場合も最も高い類似度を示すのは B, C の結果と同様であるが、Isowighteone に対する類似度の順位が逆転していることがわかる。サポートを高くする

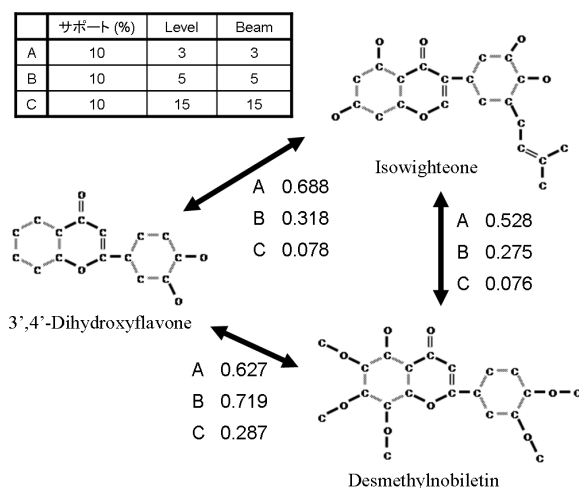


図 7 パラメータ変化と類似度 1

Fig. 7 Variations of Similarity values, # 1

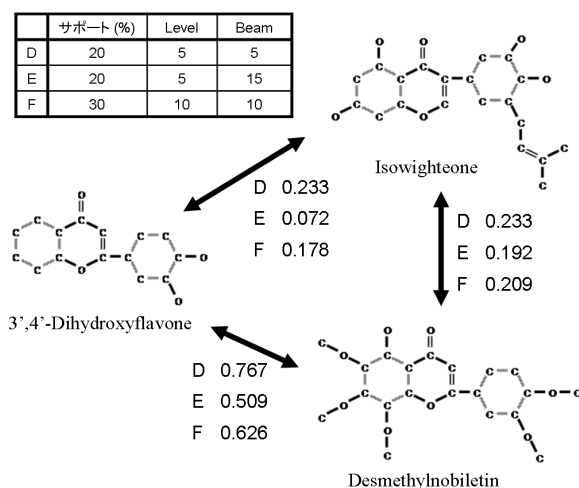


図 8 パラメータ変化と類似度 2

Fig. 8 Variations of Similarity values, # 2

ことで、大きな部分構造の持つ影響力が弱くなることで、差分構造の類似性が高く評価されているのではないかと推測できる。

5.4 特徴抽出システムを用いた知識発見

フラボノイド類の化学物質に対して、SDA クラスタリング法を適用し、結果について考察と結果の活用方法を示す。使用する部分構造は、CI-GBI 法のパラメータを (10, 10, 10) で抽出される 483 種類の部分構造を用いる。クラスタリング手法は、非階層的クラスタリングである k-medoid 法を用いて、28 個のクラスに分割した。生成されたクラス内に含まれる物質の一例を図 9 に示す。また、クラスタリングを行った結果の各クラス事例数と共通構造のイメージを図 10 に示す。図 10 に示す共通構造によって、図 9 に示されているようなクラス内のグラフの特徴は、容易に捉えることが可能である。また、この共通構造同士を比較することでクラス間の特徴の違いを視覚的に捉えることが可能である。

次に、Cluster 22 の共通構造と差分構造を図 11 に示す。共通

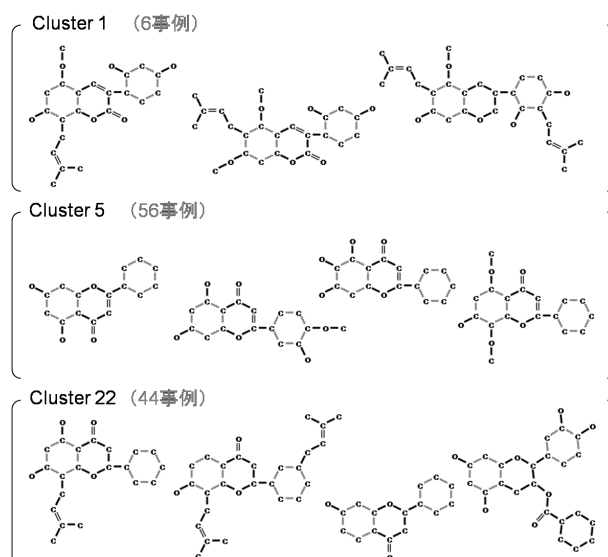


図 9 生成クラスと含まれる物質の例

Fig. 9 Examples of extracted clusters in flavonoids data

クラス#	Cluster 1	Cluster 2	Cluster 3	Cluster 4
事例数	6	4	9	14
共通構造				

クラス#	Cluster 5	Cluster 6	Cluster 7	Cluster 8
事例数	56	5	11	5
共通構造				

クラス#	Cluster 9	Cluster 10	Cluster 11	Cluster 12
事例数	3	5	5	8
共通構造				

図 10 クラスタリング結果と共通構造

Fig. 10 Clustering results and maximum common graphs

構造の各ノードは、一意に識別できるようにラベルを振りなおしたものである。ただし、N1 や N2 のように対象になっているノードは同じラベルが割り当てられている。差分構造においては、例えば、part11 と part13 は同じ構造であるが、それぞれ共通構造の N3 と N11 の異なる位置のノードに接続されていることから別の構造とみなしている。また、ここで得られた差分構造は、元々の物質に含まれていたかという情報が付加されている。その一例を図 12 に示す。この情報を参照することで、化学物質の特性値などから抗菌活性値の低い物質に含まれる構造の発見を容易にし、化学物質の物理化学的性質と構造情報を結びつけた分析が可能となると考えられる。

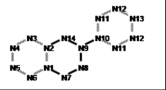
差分構造 イメージ					
差分構造名	part 001	part 002	part 003	part 004	part 005
含有数	5	19	15	13	26
差分構造 イメージ					
差分構造名	part 006	part 007	part 008	part 009	part 010
含有数	3	4	1	1	23
差分構造 イメージ					
差分構造名	part 011	part 012	part 013	part 014	part 015
含有数	5	1	1	1	4
差分構造 イメージ					
差分構造名	part 016	part 017	part 018	part 019	part 020
含有数	1	1	4	2	5
差分構造 イメージ					

図 11 Cluster 22 の最大共通構造と差分構造
Fig. 11 Maximum common graph and difference-subgraphs in Cluster 22

差分構造part 001	含有グラフ名	抗菌活性値
	Taxifolin	256
	Catechin	256
	Epicatechin	256
	Epigallocatechin	64
	Gallocatechin	64
差分構造part 011	含有グラフ名	抗菌活性値
	Glabranin	32
	Glabrol	4
	Kushenol W	64
	Clepidotin B	64
	3-Hydroxyglabrol	16
差分構造part 013	含有グラフ名	抗菌活性値
	Kazinol B	20

図 12 差分構造と物質の抗菌活性値の一例
Fig. 12 Difference-subgraphs indicating MIC values derived from flavonoids including them

6. まとめ

本稿で提案した特徴抽出システムは、対象とするグラフ集合の特徴を反映した構造類似性を SDA 法を活用して実現したものである。特に、それに基づくクラスタリング結果の活用として、各クラスタの共通構造と差分構造を提示することより、クラスタの理解容易性は飛躍的に向上したといえる。その結果、有用な知識の獲得が期待される。

また、SDA 法により求められる構造類似性が、用いる部分構造群によってどのような影響を受けるのかを検討した。その結果、一定以上の部分構造を抽出した状態であるなら、構造的特徴を反映した安定的な類似度を提供することができることが確認できた。ただし、これについては、厳密な見解を示すに至っておらず、構造分布と類似度の関係性、部分構造群の抽出基準の明確化など、より詳細な検討が必要である。

今後の課題としては、システムのより実的な活用に向けた取り組みとしては、まず、化学物質を対象とした新規化合物生成へのより有効な情報の提供を実現することがあげられる。さらに、化学構造データ以外のグラフ構造データに対しても適用事例を積み重ね、より汎用性の高いシステムとして展開していきたい。

文献

- [1] 速水亜希子, 稲積宏誠: 部分構造の包含関係を指標とするグラフクラスタリングの提案—化学物質を対象として—, SIG-KBS-A405, pp.1-6 (2005)
- [2] Kuramochi, M. and Karypis, G.: An Efficient Algorithm for Discovering Frequent Subgraphs, *IEEE Trans. Knowledge and Data Engineering*, Vol.16, No.9, pp.1038-1051 (2004)
- [3] 松田喬, 元田浩, 鷲尾隆: 一般グラフ構造データに対する Graph-Based Induction とその応用, 人工知能学会誌, Vol.16, No.4, pp.363-374 (2001)
- [4] Matsuda, T., Motoda, H., Yoshida, T. and Washio, T.: Mining Patterns from Structured Data by Beam-wise Graph-Based Induction, *Proc. of DS2002*, pp.422-429 (2002)
- [5] Nguyen, P., Ohara, K., Motoda, H. and Washio, T.: Cl-GBI: A Novel Approach for Extracting Typical Patterns from Graph-Structured Data., *Proc. of PAKDD2005*, pp.639-649 (2005)
- [6] Palmer, C., Gibbons, P. and Faloutsos, C.: ANF: A Fast and Scalable Tool for Data Mining in Massive Graphs, *Proc. of the KDD-2002*, pp.81-90 (2002)
- [7] Takahashi, Y., Ohoka, H. and Ishiyama, Y.: Structural Similarity Analysis based on Topological Fragment Spectra, *Advances in Molecular Similarity*, Vol.2, pp.93-104 (1998)
- [8] 和田貴久, 大野博之, 稲積宏誠: 対象グラフ集合の特性を反映した構造類似性の提案, 日本データベース学会 Letters (DBSJ Letters), Vol.6, No.1, pp.185-188 (2007)

和田 貴久 Takahisa WADA

2008 年青山学院大学大学院理工学研究科博士前期課程修了。人工知能学会学生会員。日本データベース学会学生会員。

大野 博之 Hiroyuki OONO

青山学院大学理工学部情報テクノロジー学科助手。2002 青山学院大学大学院理工学研究科修了。情報処理学会学生会員。

稲積 宏誠 Hiroshige INAZUMI

青山学院大学理工学部情報テクノロジー学科教授。1984 早稲田大学大学院理工学研究科博士前期課程修了。工学博士。日本データベース学会、電子情報通信学会、情報処理学会各会員。