

スーパーコンピュータを利用した 自治体例規の形式的分析

角 田 篤 泰
齋 藤 大 地
大 谷 忠

- 第1章 はじめに
- 第2章 スーパーコンピュータを利用した類似度計算
- 第3章 集計方法
 - 第1節 例規データ
 - 第2節 グループ化の手法
- 第4章 集計結果
- 第5章 結果分析
 - 第1節 集計結果の概観
 - 第2節 集計結果についての考察
- 第6章 まとめ

第1章 はじめに

本稿は全国802自治体の公開されていた全例規について、文字の並び方の観点からの類似度をスーパーコンピュータによって形式的に算出し、その情報を基にした分類結果について報告するものである。処理方法の詳細については、前号にて発表済みであるので¹⁾、本稿では省略させて頂く。本稿は、その処理の結果をより詳細に分析したものとして位置づけられる。

本稿で示す分析結果とは、例規の形式的な類似度を基に、似ている例規ごとにグループ化して分類した時、そのグループの分布状況はどのよ

うになっているかを示すものである。一言で表現できるような分布状況ではないが、客観的な数値として示すことができるので、今後の様々な研究の拠り所になると期待している。さらに、筆者らが目指しているものは、単に計算結果を整理して示すことだけではなく、本研究を含むe-Legislationの研究プロジェクト²⁾の中でも活用できるような情報を抽出することである。具体例としては、例規のデータベースなどで検索された膨大な例規群を自動的に分類させたり、準則のような例規のテンプレートとなる雛形を自動合成するような機能に生かしたりすることを想定している。そのためには、典型的なパターンの抽出が必要になる。典型的なパターンとは、多くの同様の制度の記述に現れる構造や表現である。本来は意味内容で考えるものであるが、外形的にもそのようなパターンは現れるものであるから、そこに着目するのである。そのようなパターンは同じような記述として条文上に現れるので、本研究のような文字列の並びに基づく類似度の情報があれば、ある程度は獲得可能となる。そこで、本稿では、集計結果として、どのような規模のグループがどの程度存在しているかという一般的な整理の他に、パターン抽出に役立つようなグループの存在可能性を示したいと考えている。これが本稿の目的である。

このような意図で分析を行った結果、1例規につき1グループにしか所属できないような厳格な方式で分類すると、有益なグループは抽出できなかったが、グループ間で例規が重複することを許すような分類方式でグループ化を行うと、多くのグループ化が実現できることが示された。これによって、今後、例規の典型的パターンの提示に貢献できると期待される。なお、既に名古屋大学大学院法学研究科附属法情報研究センターから全国の自治体に向けてリリースしていた旧条例データベースの内部では、このようにして得られた分類情報が生かされていた。

本研究では、例規の形式的側面、すなわち字面に焦点を当てているが、もちろん、これは意味内容を軽視するものではない。もともとe-Legislationプロジェクトでは、法令の意味の部分の可視化を行うことが大きなテーマである。しかしながら、そのためにも機械的にできるような前処理があれば、先に済ませておきたいのである。いくら複雑であっても本質的に機械的な処理であれば、人間の労力を軽減させて、その分

の労力は意味内容に関する部分やより人間的な部分に向けたいと考えているのである。そのための形式的なアプローチであることをお断りしておく。

最後に、本稿の構成であるが、まず、第2章で今回使用するスーパーコンピュータの計算結果がいかにして算出されたかについて簡単に概要を示し、次に第3章で今回の分析の方法、第4章でその結果を示し、最後に、第5章で考察内容を示す。第6章は全体の簡単なまとめと今後の課題を記している。なお、付録として、今回分析した例規の収集対象となった自治体の一覧を提示した。

第2章 スーパーコンピュータを利用した類似度計算

本章では、今回の分析対象である例規類似度のデータを算出した、スーパーコンピュータを用いた例規類似度の計算処理について概要を示す。

まず、簡単に処理手続きのフローを示すと、①例規をテキストデータ形式で収集する、②計算しやすいように加工する、③条文単位で、全ての条文の組合せについて形式的類似度を計算する、④算出された条文単位の類似度情報を基に例規単位の類似度を算出する、というものである。この中で、スーパーコンピュータを用いたのは③の処理である。

この③の処理における形式的類似度とは、「編集距離」という概念を使った文字の並びの観点から計測した類似度である³⁾。簡単に言えば、ある文Aを別の文Bに書き換える時、挿入や削除などの最小の編集手続きを何回行ったかを計測することによって、AとBがどれくらい似ているかを評価する尺度とするものである。この場合、似ていなければ多くの書き換えが必要になるため、書き換え回数を一種の「距離」とみると、その「距離」が増えるため、「AとBは遠い」と考えるのである。逆に、ほとんど書き換えがない場合は、回数、すなわち「距離」が減り、「AとBは近い (=似ている)」と考えるのである。この手法は、情報科学系の言語処理の分野や古くは通信方式の分野などでも非常にポピュラーなものである。ただし、この計算コストは高く、何の工夫もなく、そのままこの処理を実施してしまうと、今回のデータ規模の場合なら、通常のパソコンなどでは数百年もかかってしまう。なんとかアルゴリズム

を工夫しても、半年程度かかってしまうため、本研究ではスーパーコンピュータを用いることにした。なお、実際の計算時間は、処理待ち時間やデータの入れ替え時間などを含めても、4日程度であった。

このような尺度によって類似度を計測した後に、どの程度の類似度であれば「似ている」と判断するか、と言う点が問題となる。本研究では、これを2段階で判定している。まず、条文と条文との類似度は、全く同じ条文で書き換えのない場合を10割似ているとした時、3割以内の書き換えで済むようなケースを類似した条文同士とみなしている。これが先の③の計算に基づく判定結果として出力される。さらに、例規と例規の類似度については、このような類似条文が全体の7割以上を占めているような例規同士を互いに類似している判定している。これが先の④の計算に基づくものである。なお、③の判定には膨大な計算コストがかかるため、スーパーコンピュータを用いて計算しているが、④はパソコンでも数日程度の計算時間で実施可能なため、0割（正確には1条だけ類似条文を含むケース）から10割までの11段階のケースについて算出している。本稿では、これらの各ケースについても計算結果の概要を示す。

本稿では、以上のようにして得られた計算結果を基に、さらに分析した結果を示す。その方法や結果については次章以降で示す。

第3章 集計方法

第1節 例規データ

本節では、スーパーコンピュータでの計算時に入力された、本分析の対象となる例規データについて確認しておく。前述の通り、これらは802自治体の全例規である（各自治体名は本稿の付録を参照）。ただし、附則と別表は除いている。これらは2009年5月に収集したものである。なお、今回の集計に関連した数値情報は次の通りである。

総例規数……………	455372本
総条文数……………	5595807条
1自治体あたりの平均例規数 ……	約568本（＝総例規数/802）

これらの例規データから類似度判定する方法については前章で記した通りであり、その判定結果のデータが今回の分析対象である。なお、例規内の7割の条文が類似であることを持って類似と判定して、様々な分析を行っているが、本稿では、念のため、3割～10割のケースについても、類似例規の分布度合いを示しておく。ここで、0割～2割のケースは計算量が膨大であるため、これ自体にもスーパーコンピュータを利用したいくらいの規模であるのに対して、そのケースは、例規のほんの一部に似た条文がある、という状況を示すだけのものであるから、特に実証しなくても、ほぼ当然のことであると思われるので、割愛している。

第2節 グループ化の手法

本節ではグループ化を行うための手法について述べる。今回の集計作業では、次の2つの方式を用いた。

- 方式1) 同値関係としての類似性に基づくグループ化
- 方式2) 例規ごとに類似例規を割り当てるグループ化

方式1ではグループ化された後のグループについては、そのグループ内の任意の2つの要素についても必ず似ている、という性質が成り立ち、グループ間で重複するような要素を持たないという性質を持つ。すなわち、ある例規が1つのグループに分類されていたら、他のグループにも分類されているようなことはない、という厳しい制約である。これに対し、方式2では、グループ作成時に中心となった1つの例規に似てさえいれば、グループ内例規間では似ていると判定されなくても構わないような緩い制約でグループを形成して分類したものである。

それぞれについて、順に詳しく説明する。

(1) 同値関係としての類似性に基づくグループ化

同値関係とは、直観的には「=」の記号で表されるような2者の関係性のことであり、「等しい」というニュアンスのある関係性を指す。このような関係性は数学的に厳密に定義され、その性質も十分明らか

型例の取りこぼしを防ぎたいと考えたからである。特に、(1)の方式では小規模なグループしか獲得できなかったため、逆に広めにグループ化を行いたかったのである。

なお、この方式を採用することによって、集計結果は、自動的に次の性質を持つことになる。すなわち、全グループ数は、全体集合の要素数と一致する、というものである。

こうして両者を見比べると、概念の分類などを行う場合は、前者のようなきつい制約がなければ、きちんと分類できないが、我々のような大きな枠組みとしての典型パターンを探索する道具立てとしては、必ずしも前者のような厳密さに縛られる必要はなさそうである。

ここで、両方式のトレードオフについて整理しておく。(1)の方式の利点は、グループ間で要素の重複がなく、はっきり別のグループとして区別できることやグループ内での任意の2者が類似していることが保障されるような厳格なグループ化になっていることであり、欠点は、任意の2者の類似性の保持(推移則の成立)の要請が厳しいため、一般に、大きなグループを獲得したり、獲得グループ数を多くしたりすることが困難なことである。

一方、(2)の方式は、欠点としては、グループ間で要素の重複があるので、はっきり分類されているとは限らず、グループ間の関係が曖昧となりがちである。一方、長所としては、中心となる要素に似ているものであれば良い、という緩い制約のため、要素を広く収集することができる。その結果、大きなグループや数多くのグループを獲得しやすくなる。

今回のグループ化の方式は、いずれも必ずしもベストである訳ではない。応用方法にも依存するし、もっと例規内部の意味的側面や付属情報を利用して、より現実的で意義のあるグループ化を行うことも重要である。実際、コンピュータを用いたグループ化(「クラスタリング」や「分類」でもある)の方式はいくつも存在する。ただし、データの分類方式の多くは、直接対象データを分析して分類する方式が主流である⁵⁾。例えば、最も類似しているペアのグループから、徐々に類似グループを結合して、各分類グループを広げてゆく階層的クラスタリングの方式がある。あるいは、K平均法のようなものもあり、先にグループのサイズや個数を決

めておいて、そのグループに各データ結び付けてゆく方式である⁶⁾。

これらの手法は、本研究の場合ならば、例規の記述内容を直接分析して分類する必要がある。しかしながら、今回の実験では、既に類似性の情報を持ったデータを保持していて、むしろ、データそのものの属性情報はなくなってしまっている。そのため、従来の方式が単純に適用できないのである。もちろん、工夫次第で、類似性の情報を属性情報に読み替えることも可能なので、今後の検討課題ではある。

この他にも、模範的な分類例（正解の例）をいくつか与えておいて、機械学習を行いながらクラスタリングを実施する手法も様々なものが提案されており⁷⁾、成果も上がっているが、本研究の場合には、一体何が正解となるのか、それ自体が分かっていないので、現段階では導入することができない。

第4章 集計結果

まず、集計結果のうち、例規類似度7割の場合の数値データの一覧を示す。なお、本章では、各グループ内に所属する例規数をそのグループの「サイズ」と表記することにする。

[A] 同値関係としてのグループ化

全グループ数（＝サイズ3以上のグループ総数）	5126
非グループ数（＝サイズ1とサイズ2のグループ合計）	421739
他に類似例規のない例規数（＝サイズ1のグループ数）	409950
最大グループのサイズ	38
全グループ（サイズ3以上のグループ）の延べ例規数	21844
他に類似例規のない例規の例規全体に対する割合	90%
グループ（サイズ3以上のグループ）の平均サイズ	4.3
グループ（サイズ3以上のグループ）サイズの中央値	3

[B] 例規ごとに割り当てるグループ化

全グループ数（＝サイズ3以上のグループ総数）	40896
非グループ数（＝サイズ1とサイズ2のグループ合計）	414476

他に類似例規のない例規数 (=サイズ1のグループ数) ……	396877
最大グループのサイズ……………	1770
全グループ (サイズ3以上のグループ) の延べ例規数 ……	893143
他に類似例規のない例規の例規全体に対する割合……………	87%
グループ (サイズ3以上のグループ) の平均サイズ ……	21.8
グループ (サイズ3以上のグループ) サイズの中央値 ……	7
大型グループ数 (=サイズ200以上のグループ数) ……	558
超大型グループ数 (=サイズ800以上のグループ数) ……	45
サイズ2以上の全グループの延べ例規数……………	928341

次に、例規類似度7割の集計結果の度数分布表(表1)とヒストグラム(図1、図2)を示す。度数分布表ではグループサイズ(グループ内例規数)が階級であり、度数が0の階級は省いている。ヒストグラムについては、図1の場合、サイズが1や2のものは、グループとは言えない上、桁違いに数が多く、グラフが見難くなってしまうので、割愛している。また、図2の場合も、サイズが50未満の小さなグループについては、典型パターンとしての意味合いが薄い上、桁違いに数が多く、グラフが見難くなってしまうので、割愛している。

最後に参考のため、例規類似度3割~10割の分布状況について、同値関係としてのグループ化の方式で集計したヒストグラムを示す。なお、サイズが15未満の小さなクラスはどの割合でも桁違いに数が多く、一方、サイズが100以上のクラスはほとんど存在しないことから、これらを入れるとグラフが見難くなるため、割愛した。

なお、紙面の関係で今回報告できなかった詳細の集計結果についても、今後、名古屋大学大学院法学研究科附属法情報研究センターのWebサイトから公開する予定である⁸⁾。

〈62〉 スーパーコンピュータを利用した自治体例規の形式的分析（角田・齋藤・大谷）

表1：例規類似度7割度数分布表

同値関係方式		例規割当て方式	
グループ内例規数	個数	グループ内例規数	個数
1	409950	1～49	452070
2	11789	50～	1777
3	2781	100～	694
4	1025	150～	273
5	509	200～	100
6	277	250～	94
7	175	300～	78
8	101	350～	57
9	79	400～	25
10	44	450～	44
11	27	500～	31
12	23	550～	28
13	20	600～	21
14	20	650～	14
15	11	700～	14
16	9	750～	7
17	5	800～	5
18	2	850～	7
19	2	900～	5
20	1	950～	6
21	1	1000～	5
22	2	1050～	4
23	1	1100～	2
24	2	1150～	5
26	2	1200～	1
27	3	1250～	0
30	1	1300～	1
35	1	1350～	2
37	1	1400～	1
38	1	1450～	1

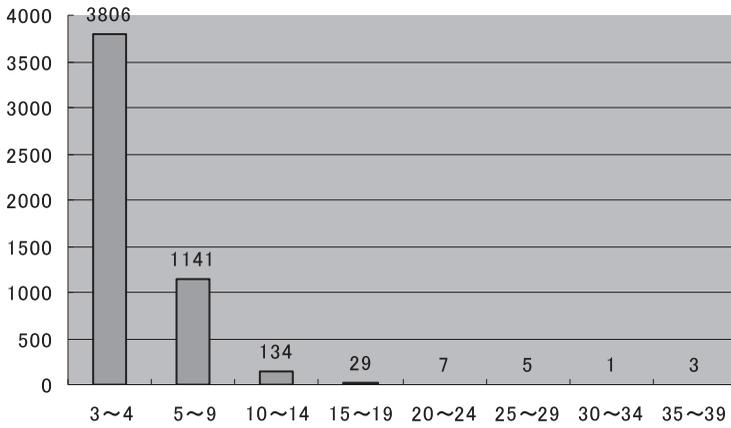


図1：例規類似度7割のヒストグラム（同値関係方式）

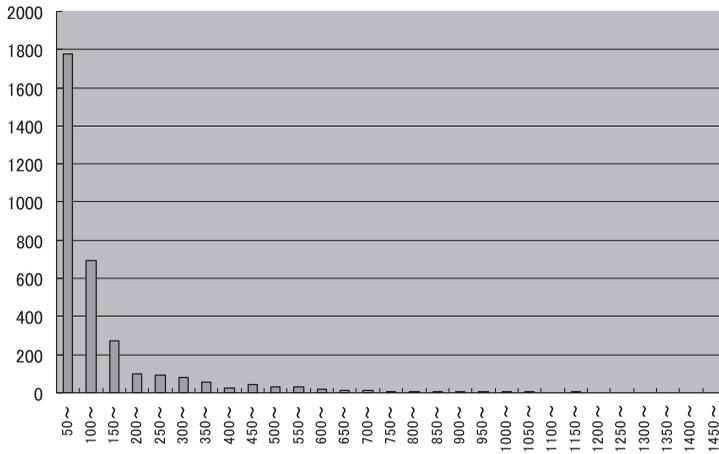


図2：例規類似度7割のヒストグラム（例規割当て方式）

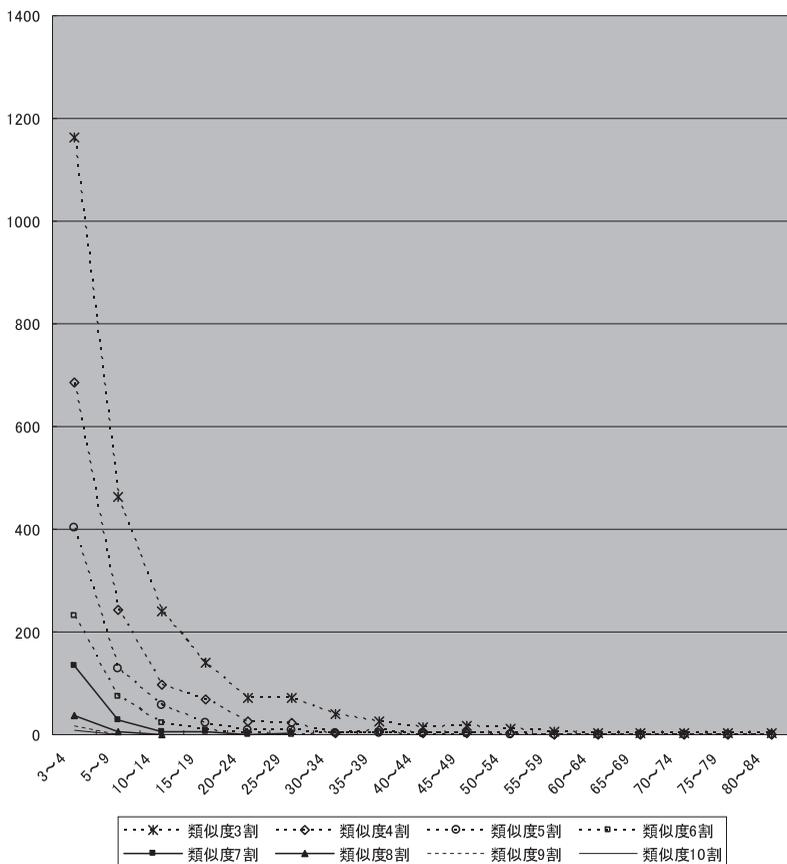


図3：例規類似度3割～10割のヒストグラム（同値関係方式）

第5章 結果分析

第1節 集計結果の概観

1. 例規類似度7割で同値関係に基づくグループ化のケース

前章で示されたデータを概観すると、最大グループでも、そのサイズは38であり、典型例となるようなグループを構成している訳ではなく、

こちらの方式ではテンプレートとなるような共通の例規パターンはほとんど得られないことになる。

この最大グループは、小・中学校の設置に関するものであり⁹⁾、ついで、サイズ37のグループが地域振興基金に関するもの¹⁰⁾、サイズ35の施設維持や公用地などに関するもの¹¹⁾、サイズ30の教育機関等の長に関するもの¹²⁾、と続く。

なお、この集計によって、7割のケースで、他に類似例規がないと判断されたものの中には、例規名だけ見ると、「〇〇市会計規則」となっており、同じような記述であろうと思われるものがあるが、実際に、各条文を見比べてみると意味内容は同じものも含まれるが、表現や構成が大きく異なるものがある¹³⁾。また、グループの形成には至っていないが、非常に似ているペアを形成している組、すなわち、サイズ2のグループ(ペア)の中には、札幌市と函館市の選挙運動資金の公的負担の条例¹⁴⁾や、甲府市と大月市の同様の条例¹⁵⁾のように、地域的な参照関係あるいは協力関係を連想させるものもある。この傾向はサイズ3でも同様で、例えば、北海道の赤平市、士別市、名寄市はこの同値関係のきつい縛りでグループ化しても、同じグループとみなされるような健康保険税条例や健康保険条例を持っている。

2. 例規類似度7割で例規ごとに割り当てるグループ化のケース

前章で示されたデータを概観すると、サイズ200以上の大きなグループが558件あり、典型例となるグループを構成しているものが含まれる可能性が高い。この集計結果では、最大の例規数のグループは1770本の例規数であり、このグループの例規は高額療養費貸付や特別養護老人ホームに関するものであった¹⁶⁾。また、次に数が多かったのは1443本であった。これらの例規は教育委員会に関するものであった¹⁷⁾。

これらのグループには800本以上もグループ内に例規を含むものが45グループもあったが、今回の調査対象の自治体数が802団体であるから、単純に考えると、ほぼ全ての自治体で採用されているような例規が最大で45パターン存在する可能性が示唆される。もし、これらの45グループの間にあまり重複がないとすれば、自治体の平均例規数は約568本であることから、どの自治体でも、その約1割弱の例規が全国に共通する

内容を持っている可能性もある。なお、サイズが802以上のグループであれば、当然であるが、1つの自治体で2つ以上の例規をそのグループ内に所属させていることになる。つまり、同一自治体内でも類似例規が存在している、とすることである。例えば、800本に近かった791本の類似例規を含むグループでは、出現する自治体は331団体であるので、平均すると、それらの自治体はこのグループの例規に類似する例規を約2.4本持っていることになる。

3. 例規類似度7割以外の同値関係に基づくグループ化のケース(参考)

前章の図3の例規類似度3割～10割のヒストグラム（同値関係としてのグループ化）によれば、どの割合でも度数分布状態の形は似ており、サイズの小さなグループが圧倒的に多数であり、サイズが大きくなると、急激にその数は減ると言うものである。この点においては、7割のケースが特別に変わった分布ではなく、着目したサンプルとしても、問題がないと考えられる。

10割という、きつい縛りの中での最大グループはサイズ15という小規模なものであり、教育機関やその長に関する例規が中心であった¹⁸⁾。また、8割のケースでも、最大グループ（サイズ34）では、教育機関やその長に関する例規が中心である¹⁹⁾。これは、前述のように7割のケースでも同様であった。ただし、細かく見ると、教育機関の長に関するものは4位となっていて、小中学校などの設置条例が最大グループであった。

さらに類似度を下げて6割にすると、最大グループ（サイズ119）の例規は様々な給付金に関するものであった²⁰⁾。ちなみに、3割のケースの最大グループは、大分大きなものとなり、408本の例規が含まれ、そこには、教育委員会公告式に関する規則が多く見受けられる²¹⁾。

第2節 集計結果についての考察

本稿の目的は、例規の組合せの類似度を基にグループ化（集計）を試み、例規の分布状態の概要を示すとともに、例規のテンプレートとなるような典型的なパターンが、どれくらいありそうか示すことにある。分

布状態については、前章および前節に示したが、典型的なパターンについての考察は本節で行う。なお、本節では、例規類似度7割のケースであることを前提として考察を進めている。

先の集計結果より、次のようなことが結論付けられる。

- ①同値関係に基づくグループ化では典型パターンが抽出できない。すなわち、大規模な類似例規のグループは存在しない。
- ②例規ごとに割り当てるグループ化では、ある程度の典型パターンの存在が暗示されている。ただし、重複があるので、厳密にはそのパターン数は特定できず、そのパターンは最低で35程度から最高で600程度の範囲と予測される。

それぞれについて、以下、順に記す。

(1) 同値関係に基づくグループ化

この場合、最大グループでも38本であるので、自治体数の802から考えると、1/20弱であり、どこの自治体にも存在する典型的な内容を持った例規のパターンという程のものではない。そして、サイズ3以上のグループの平均サイズは4.2（＝サイズ3以上のグループの延べ例規数/サイズ3以上のグループ数）であり、サイズ10以上のものでも、179個に留まるので、何らかのパターンなのかも知れないが、広範に使われるものは少ない。少なくとも典型パターンと呼ばれるような大規模なグループは存在しない。

なお、同値関係に基づくグループ化の結果は、客観的なものであるもので、一応受け入れるとしても、我々が経験上得ている直観とは異なる。つまり、準則や準則集のサンプルをベースに例規を作成しているケースが多いはずであり、少なくとも従来の蓄積から、そのような例規が多数存在すると考えられる。その場合、かなり多くの自治体が同じような構成と内容で条文を記述しているはずである。それにも関わらず、典型例のような類似例規のグループが形成されていないことに疑問を感じるのである。多分、その理由は、今回の例規類似性の基準が同値類の考え方に倣った、厳格なものなので、準則などの基となる

テンプレートが同じでも、その展開の仕方まで同様でないと、同値とはみなされないことにあると考えられる。すなわち、この方式のグループ化の要請がきついのである。なお、実際今回の結果を受けて、筆者らが全国の自治体に提供する条例データベースでは²²⁾、初期には同値関係に基づく分類方式を利用して、類似例規の自動分類を行っていたが、あまり役立たないため、この方式を止めて、別の分類方式を一時的に採用していた。

（2）例規ごとに割り当てるグループ化

同値関係に基づくグループ化の際の条件がきついことを受けて、条件を緩和した方式でグループ化を試みたものが、例規ごとに割り当てるグループ化の方式である。この結果、200個以上類似している例規が存在するようなグループは558件ある。これらはテンプレートとなるような典型的例規であり、200以上の自治体が参照しているケースが558件あるということでもある。すなわち、558の典型パターンが存在する可能性を示唆していることになる。ただし、同値関係に基づく集計方法とは異なり、複数グループに同一例規を含むような重複を認める方式なので、558件のうちには、同様のものが含まれてしまっている可能性があるため、最大で558件であり、実際にはより少ない件数と予想される。ここで、もう少し細かく見積もることにする。まず、何らかのグループに所属する例規数は、何も所属しない例規数（＝サイズ1のグループ数）を全例規数から減じたものなので、58495本である。次に、サイズ2以上の全グループについて、それらに含まれる例規の延べ数は計928341本であり、単純に平均すると、1本の例規当たり約16のグループに出現することになる。もし例規ごとに1つの観点でグループ化されるなら、グループ数は16倍に膨れ上がっていることになるため、 $558/16 = \text{約}35$ となり、約35のパターンが存在していると予測される。しかしながら、そもそも観点は多様なので、同じ例規に対しても、別の観点で類似性が捉えられるならば、重複が存在しても自然なことである。むしろ、全く重複のない、同値関係に基づくようなきつい類似性の方が、不自然とも言えるであろう。実際、前述のように、パターン抽出の方式としては、あまり役立たなかった。

重複があっても1つの例規が複数の観点や多くの部分から成り立つ以上、それらの観点から見た類似性や各部分の共通性が捉えられれば、複数のグループに割り当てられることも当然ありうる。

このようにして、必ずしも、重複が問題を孕むようなものであるとは限らず、重複は自然なことで考えられる。すると、典型パターンが約35だけである可能性は低くて、もう少し有効な典型パターンは多く含まれていると考えられる。

さらに、重要なことは、本稿が目指しているのは、元々より広く用いられるような例規のパターンを見つけることにある。このような場合、広くグループ化を行う方式の方が、有用性が高い。同値関係の方式のように、いくら細部まで類似しているものを検出できても、最終的に計上されるグループ内の例規が少ないのでは、広く使われる典型パターンとは言いがたい。このような理由から、例規ごとに類似例規を割り当てる方式の方が、実験結果としても、有効であり、一定程度の典型パターンの存在可能性を示したと言える。

最後に、「パターン」という用語について付記しておく。本稿では厳密に区別して用いていなかったが、ソフトウェア工学の分野では、実は「パターン」と言う語を「フレームワーク」と言う語から区別して論じることがある。前者は、様々な局面に詳細な部分品として現れるような汎用な型のことであり、後者は様々な局面で大きな雛形として、部分が無視(=匿名化=抽象化=変数化)して、見出されるような型のことを指す。本研究、とりわけ、本稿で検出しようとしていたパターンとは、実はこの用語法に従えば、むしろ「フレームワーク」である。一方、もし、この用語法における例規の「パターン」を抽出するのであれば、それは、あまり重要なことではないかも知れない。なぜなら、細部の書き方はほとんどの場合、法制執務の記述ルールに則っているため、頻出するのは当然であるし、既に多くの「パターン」を記したマニュアル類も多数存在しているからである。また、既存の法制執務の支援ツールの中にはそのような用語をある程度自動でチェックしてくれるものも存在する。さらに、筆者らが提供する例規データベース・システムの中でも、文脈検索機能として、使用例を即座に一覧にしてくれるツールを既に提供している²³⁾。このような状況

であるから、「パターン」の抽出よりも「フレームワーク」の抽出の方が現実の貢献度が高くなるものと考えられる。そうなると、やはり、厳しく一致を求める同値関係に基づく方式よりも、例規ごとに他の類似例規を集める方式の有効性の方が高いと言えるであろう。

第6章 まとめ

本稿では、802自治体の例規を分析した実験結果として、類似例規のグループの分布状況を示した。その際に、同値関係が成り立つような似たもの同士でグループ化を行う厳密な方式では、最高でも38個の例規から成るグループが形成されるだけであった。この方式では典型的なグループ、すなわち、広範に利用される典型パターンのようなものは検出できなかったことになる。しかしながら、グループ化の基準を緩めて、グループ間での重複を認めた場合には、最低でも35程度以上の典型パターンが存在する可能性を示すことができた。現時点では、存在可能性を示したに過ぎないので、今後の課題は、典型パターンを具体的に特定して検証を進めることである。

本研究プロジェクト自体は、今後も続くものであり、特に、全国の自治体に直ちに現場で利用できる形で、条例データベースを提供する必要があるため、対象としている例規をなるべく最新のものに更新する必要がある。その際に、今回の計算に用いたデータの収集から3年を経ているので、最新のものを収集して、再計算と再分析を行うつもりである。これによって、この3年間でどのようなタイプの例規が増えているのかということも示されるはずである。まずは、このような作業を行うことが当面の課題である。また、本研究は字面上の形式的分析に焦点を当てているが、言葉は意味を表すものであるから、語句の意味的な観点からの分類を行ったり、統計的分析を進めたりすることも今後の課題である。

最後に、本稿で行った分析作業の位置づけは、より大きな観点から見ると、第1章で記したようにe-Legislationの一環とみなすことができるが、今回のような類似性に関する問題意識としては、そもそも筆者が進めてきた大きなテーマに含まれるものである²⁴⁾。その観点からは、文脈に応じて切り替わるような動的な類似性の抽出方法が基本的関心とな

る。今回はまだそのような本来の問題関心に基づく考察や開発が行われていないので、今後はそのような点についても研究・開発を進めてゆきたい。

謝辞

本稿執筆作業において、名古屋大学大学院法学研究科研究員の久田亜有美氏の協力を得た。この場を借りて感謝の意を表したい。なお、本報告における分析計算やデータ収集では、JSPSの科研費基盤研究（A）課題「e-Legislationに基づく法制執務方法論の情報科学的基礎付けと検証」（24240040）の助成を受けている。

付録 例規収集を実施した自治体一覧

北海道	赤平市 厚真町 網走市 池田町 今金町 岩見沢市 雨竜町 枝幸町 恵庭市 遠軽町 雄武町 大空町 置戸町 音威子府村 音更町 夕張市 北広島市 清里町 倶知安町 札幌市 更別村 士別市 島牧村 白老町 寿都町 大樹町 苫前町 豊浦町 長沼町 中富良野町 名寄市 函館市 羽幌町 浜頓別町 日高町 広尾町 深川市 北斗市 三笠市 南富良野町 芽室町 紋別市 八雲町
青森県	青森市 板柳町 今別町 おいらせ町 大鰐町 七戸町 外ヶ浜町 田子町 つがる市 十和田市 南部町 野辺地町 平川市 深浦町 藤崎町 三沢市 むつ市 六戸町
岩手県	岩手町 大槌町 金ヶ崎町 北上市 久慈市 九戸村 平泉町 洋野町 藤沢町 矢巾町 陸前高田市 釜石市
宮城県	石巻市 岩沼市 大河原町 大崎市 大郷町 大衡村 加美町 栗原市 蔵王町 塩竈市 七ヶ宿町 柴田町 白石市 仙台市 大和町 富谷町 登米市 名取市 東松島市 村田町 本吉町 利府町 涌谷町
秋田県	秋田市 井川町 羽後町 大潟村 大館市 小坂町 大仙市 にかほ市 能代市 八郎潟町 東成瀬村 三種町 湯沢市 由利本荘市 横手市
山形県	飯豊町 小国町 河北町 川西町 庄内広域行政組合 白鷹町 新庄市 鶴岡市 天童市 長井市 中山町 南陽市 三川町 遊佐町 米沢市
福島県	猪苗代町 大玉村 金山町 川俣町 喜多方市 桑折町 郡山市 白河市 相馬市 棚倉町 田村市 西会津町 二本松市 磐梯町 平田村 福島県 福島市 双葉町 南会津町 南相馬市 本宮市
茨城県	阿見町 石岡市 潮来市 稲敷市 茨城県 牛久市 小美玉市 龍ヶ崎市 かすみがうら市 神栖市 北茨城市 五霞町 境町 桜川市 下妻市 常総市 城里町 大子町 筑西市 つくば市 つくばみらい市 土浦市 東海村 利根町 取手市 那珂市 行方市 坂東市 鹿嶋市 常陸太田市 常陸大宮市 ひたちなか市 鉾田市 水戸市 守谷市 八千代町 結城市
栃木県	足利市 市貝町 岩舟町 大平町 小山市 鹿沼市 上三川町 さくら市 佐野市 下野市 高根沢町 都賀町 栃木県 那須烏山市 那須塩原市 西方町 日光市 芳賀町 藤岡町 益子町 壬生町 真岡市 矢板市
群馬県	安中市 伊勢崎市 上野村 邑楽町 大泉町 神流町 甘楽町 渋川市 下仁田町 榛東村 高崎市 玉村町 千代田町 長野原町 沼田市 藤岡市 前橋市 みどり市 みなかみ町 吉岡町

埼玉県	上尾市 朝霞市 伊奈町 大利根町 小川町 桶川市 越生町 加須市 神川町 川口市 川越市 川島町 騎西町 北本市 行田市 栗橋町 鴻巣市 越谷市 幸手市 志木市 菫蒲町 白岡町 草加市 秩父市 鶴ヶ島市 ときがわ町 所沢市 戸田市 長瀨町 新座市 蓮田市 羽生市 飯能市 深谷市 富士見市 本庄市 松伏町 三郷市 美里町 皆野町 宮代町 毛呂山町 八潮市 吉川市 吉見町 和光市 鷲宮町
千葉県	旭市 栄町 いすみ市 一宮町 市原市 勝浦市 鴨川市 神崎町 匝瑳市 山武市 芝山町 大多喜町 匝瑳市 横芝光町 消防組合 袖ヶ浦市 長柄町 千葉市 銚子市 長生村 東金市 東庄町 習志野市 四街道市 野田市 船橋市 松戸市 南房総市 睦沢町 本埜村 茂原市 八千代市 横芝光町
東京都	あきる野市 足立区 荒川区 稲城市 奥多摩町 葛飾区 江東区 国分寺市 小平市 渋谷区 新宿区 墨田区 台東区 豊島区 中野区 東京都後期高齢者医療広域連合 東京二十三区清掃一部事務組合 特別区人事・厚生事務組合 羽村市 東村山市 東大和市 日野市 日の出町 檜原村 町田市 瑞穂町 港区 武蔵村山市
神奈川県	愛川町 綾瀬市 伊勢原市 川崎市 座間市 逗子市 中井町 藤沢市 南足柄市 山北町 大和市 湯河原町 横須賀市
新潟県	阿賀野市 阿賀町 出雲崎町 糸魚川市 魚沼市 小千谷市 刈羽村 佐渡市 三条市 新発田市 聖籠町 胎内市 田上町 津南町 燕市 十日町市 長岡市 新潟市 見附市 南魚沼市 妙高市 村上市 湯沢町
富山県	朝日町 射水市 上市町 黒部市 高岡市 立山町 砺波市 富山県 富山市 滑川市 南砺市 入善町 氷見市
石川県	穴水町 石川県 内灘町 加賀市 小松市 志賀町 宝達志水町 珠洲市 中能登町 白山市 津幡町 七尾市 能登町 野々市町 能美市 かほく市
福井県	あわら市 越前市 おおい町 小浜市 坂井市 鯖江市 敦賀市 珠洲市 永平寺町 五箇川公共下水道事務組合 美浜町 珠洲市 市川三郷町 高浜町
山梨県	上野原市 大月市 甲斐市 鯉沢町 甲州市 甲府市 昭和町 丹波山村 南部町 韮崎市 笛吹市 富士河口湖町 富士吉田市 南アルプス市 山中湖村 山梨県 山梨市
長野県	上松町 飯島町 飯田市 伊那市 上田市 岡谷市 駒ヶ根市 小諸市 下諏訪町 諏訪市 辰野町 千曲市 東御市 豊丘村 長和町 富士見町 松川町 松川村 箕輪町 宮田村 山形村

〈74〉 スーパーコンピュータを利用した自治体例規の形式的分析（角田・齋藤・大谷）

岐阜県	池田町 大垣市 海津市 各務原市 可児市 川辺町 岐阜市 郡上市 神戸町 坂祝町 白川町 関ヶ原町 富加町 羽島市 飛騨市 瑞穂市 御嵩町 本巣市 山県市 養老町
静岡県	熱海市 新居町 伊東市 磐田市 小山町 湖西市 御殿場市 静岡市 芝川町 島田市 裾野市 西伊豆町 袋井市 藤枝市 牧之原市 三島市 森町 吉田町
愛知県	愛西市 阿久比町 一宮市 一色町 岩倉市 大口町 大治町 岡崎市 春日井市 蟹江町 蒲郡市 北名古屋市 清須市 幸田町 江南市 小牧市 設楽町 七宝町 甚目寺町 新城市 瀬戸市 高浜市 武豊町 田原市 知多市 東海市 飛鳥村 豊明市 豊田市 豊山町 長久手町 名古屋市 日進市 半田市 美和町 弥富市
三重県	伊勢市 亀山市 紀宝町 熊野市 桑名市 玉城町 東員町 南伊勢町
滋賀県	愛荘町 近江八幡市 大津市 草津市 甲賀市 湖南市 高島市 栗東市 東近江市 日野町 米原市 守山市 野洲市
京都府	綾部市 井手町 宇治市 宇治田原町 亀岡市 京田辺市 京丹後市 京丹波町 京都市 久御山町 城陽市 精華町 長岡京市 南丹市 舞鶴市 向日市 与謝野町
大阪府	池田市 泉大津市 泉佐野市 和泉市 茨木市 大阪狭山市 貝塚市 柏原市 交野市 河内長野市 熊取町 堺市 四條畷市 島本町 摂津市 泉南市 大東市 高石市 高槻市 富田林市 寝屋川市 阪南市 枚方市 松原市 岬町 守口市
兵庫県	相生市 赤穂市 朝来市 芦屋市 伊丹市 猪名川町 小野市 加古川市 加東市 上郡町 川西市 神戸市 宝塚市 豊岡市 三木市 南あわじ市 養父市
奈良県	明日香村 斑鳩町 生駒市 宇陀市 王寺町 大淀町 香芝市 橿原市 葛城市 河合町 上牧町 広陵町 五條市 御所市 桜井市 三郷町 田原本町 奈良県 大和郡山市
和歌山県	岩出市 上富田町 御坊市 美浜町 和歌山県
鳥取県	岩美町 江府町 琴浦町 大山町 南部町 日野町 伯耆町 三朝町 湯梨浜町 若桜町
島根県	雲南市 大田市 邑南町 隠岐の島町 奥出雲町 江津市 津和野町 西ノ島町 浜田市 斐川町 益田市 松江市
岡山県	安来市 赤磐市 浅口市 笠岡市 吉備中央町 久米南町 瀬戸内市 総社市 高梁市 玉野市 庭市 美咲町 美作市 矢掛町 和気町
広島県	神石高原町 広島県 三原市 三次市

山口県	阿武町 下松市 下関市 光市 防府市 美祿市 山口市
徳島県	三好町 阿波市 神山町 鳴門市 美馬市 三好市 吉野川市
香川県	綾川町 琴平町 さぬき市 小豆島町 多度津町 三木町 三豊市
愛媛県	今治市 宇和島市 大洲市 西条市 西予市 砥部町 新居浜市 松前町 八幡浜市
高知県	安芸市 いの町 香美市 黒潮町 香南市 四万十市 宿毛市 須崎市 東洋町 中土佐町 南国市 室戸市
福岡県	朝倉市 飯塚市 うきは市 大木町 小郡市 川崎町 香春町 久留米市 黒木町 上毛町 古賀市 篠栗町 志摩町 志免町 須恵町 太宰府市 立花町 筑後市 二丈町 広川町 星野村 みやこ町 みやま市 宮若市 柳川市 八女市 吉富町
佐賀県	有田町 唐津市 佐賀県 太良町
長崎県	壱岐市 諫早市 雲仙市 大村市 佐世保市 対馬市 時津町 長崎市 長与町 松浦市
熊本県	あさぎり町 芦北町 阿蘇市 天草市 植木町 宇城市 大津町 菊陽町 玉東町 熊本市 合志市 城南町 高森町 津奈木町 和水町 西原村 山鹿市 山都町
大分県	臼杵市 大分市 珠町 国東市 佐伯市 竹田市 日田市 豊後大野市 豊後高田市 別府市 由布市
宮崎県	小林市 高鍋町 高原町 日向市 都城市
鹿児島県	出水市 指宿市 鹿児島県 鹿児島市 加治木町 霧島市 薩摩川内市 さつま町 曾於市 日置市 南九州市 南さつま市 湧水町
沖縄県	今帰仁村 本部町

注

- 1) 拙著「スーパーコンピュータを用いた自治体例規の類似度分析と例規データベースへの応用」名古屋大学法政論集246号(2012年)〈69〉-〈91〉頁。以下、「スパコン」で参照。
- 2) 拙著「e-Legislationの構想—情報処理としての立法過程—」名古屋大学法政論集241号(2011年)〈1〉-〈26〉頁。
- 3) 前掲、拙著「スパコン」。
- 4) この場合、数学的テクニクとして、どのグループにも所属しない例規は、

〈76〉 スーパーコンピュータを利用した自治体例規の形式的分析（角田・齋藤・大谷）

それ自身で要素数1のグループと考えることを前提としている。

- 5) 例えば、Toby Segaran 著、當山仁健・鴨澤眞夫訳『集合知プログラミング』（オライリージャパン、2008年）31-58頁参照。
- 6) K平均法とは、まず、ある種の中心点（グループの重心）となるものをランダムにいくつか決めておき、各データはそのいずれかの近いものに結び付けて、その中心点ごとにグループを構成する。その後、そのグループ内のデータで平均（重心）を求め、新たな中心点を再配置する。この状態で、先のグループは解消して、再び各データを近い中心点に結び付け、同様の操作を繰り返す。こうして、安定するまでこの操作を繰り返す手法がK平均法である。
- 7) 例えば、Nello Cristianini・John Shawe-Taylor 著、大北剛訳『サポートベクタマシン入門』（共立出版、2005年）では「サポートベクタマシン」と呼ばれる一種の機械学習のアルゴリズムにより、精度の高い分類を行う方式が紹介されている。
- 8) URLは次の通り。<http://jalii.law.nagoya-u.ac.jp/>
- 9) 例えば、松前町立学校設置条例（愛媛県松前町）、杵崎市立小・中学校設置条例（長崎県杵崎市）、由布市中学校の設置に関する条例（大分県由布市）など。
- 10) 例えば、白老町農業振興基金条例（北海道白老町）、田村市常葉町地域振興基金条例（福島県田村市）、島根発電用施設周辺地域振興基金条例（島根県松江市）など。
- 11) 例えば、七戸町公共用施設維持基金条例（青森県七戸町）、藤沢町まちづくり基金条例（岩手県藤沢町）、和光市公共用地取得事業基金条例（埼玉県和光市）など。
- 12) 例えば、上野原市教育委員会教育長に対する事務委任規則（山梨県上野原市）、学校その他の教育機関の長に対する事務委任規程（埼玉県吉見町）、南伊勢町教育委員会教育長に対する事務委任規則（三重県南伊勢町）など。
- 13) 例えば、貝塚市会計規則（大阪府貝塚市）と三好市会計規則（徳島県三好市）は大きく記述が異なる。
- 14) 札幌市議会議員及び札幌市長の選挙における選挙運動に要する費用の公費負担に関する条例（北海道札幌市）と函館市の議会の議員および長の選挙における選挙運動の公費負担に関する条例（北海道函館市）である。
- 15) 甲府市の議会の議員及び長の選挙における選挙運動の公費負担に関する条例（山梨県甲府市）と大月市の議会の議員及び長の選挙における選挙運動の公費負担に関する条例（山梨県大月市）である。

- 16) 例えば、益田市立特別養護老人ホーム設置及び管理に関する条例（鳥根県益田市）、三沢市高額療養費貸付基金の設置及び管理に関する条例（青森県三沢市）、白石市高額療養費貸付基金の設置及び管理に関する条例（宮城県白石市）など。
- 17) 例えば、江津市教育委員会処務規程（鳥根県江津市）、石巻市教育委員会の組織等に関する規則（宮城県石巻市）、飯豊町教育委員会会議規則（山形県飯豊町）など。
- 18) 例えば、黒潮町学校その他の教育機関の長に対する事務委任規程（高知県黒潮町）、学校その他の教育機関の長に対する事務委任規程（埼玉県加須市）、教育長に対する事務委任規則（岡山県浅口市）など。
- 19) 例えば、燕市教育委員会教育長事務委任規程（新潟県燕市）、葛城市教育委員会の権限に属する事務の一部委任及び臨時代理に関する規則（奈良県葛城市）、伯耆町教育委員会教育長の権限に属する事務の一部を学校その他の教育機関の長に委任する規程（鳥取県伯耆町）など。
- 20) 例えば、紋別市農林漁業活性化基金条例（北海道紋別市）、横芝光町介護保険給付費準備基金条例（千葉県横芝光町）、霧島市介護給付費準備基金条例（鹿児島県霧島市）など。
- 21) 例えば、白老町教育委員会公告式規則（北海道白老町）、三条市教育委員会公告式規則（新潟県三条市）、さつま町教育委員会公告式規則（鹿児島県さつま町）など。
- 22) 前掲、拙著「スパコン」参照のこと。以下のURLにも、この条例データベースの説明がある。<http://elensv.law.nagoya-u.ac.jp/project/elen/>
- 23) 拙著「e-Legislation 環境の構築へ向けて—情報科学を応用した立法過程の作業支援—」情報ネットワークローレビュー 11号（商事法務、2012年）13-32頁。
- 24) 博士論文の内容を継承した論文は、Kakuta, T. and Haraguchi, M., “A Demonstration of a Legal Reasoning System Based on Teleological Analogies”, *Proceedings of the Seventh International Conference on Artificial Intelligence and Law*, pp. 196-205, 1999. また、技術の詳細を省いた解説としては、角田篤泰・原口誠「法的推論と類似性—対話と議論の観点から」人工知能学会誌 17巻1号（2002年）14-21頁や拙著「法解釈論の知識工学的アプローチ—平井直雄「良い法律論」の形式論理的的分析と代理権濫用論への適用」北大法学論集 51巻2号（2000年）760-842頁などを参照のこと。