A Hybrid Classification Approach based on FCA and Emerging Patterns - An application for the classification of biological inhibitors

Yasmine Asses[1], Aleksey Buzmakov[1,2], Thomas Bourquard[1], Sergei O. Kuznetsov[2], and Amedeo Napoli[1]

[1] LORIA (CNRS - Inria NGE - U. de Lorraine), Vandoeuvre les Nancy, France
[2] National Research University Higher School of Economics, Moscow, Russia

Abstract. Classification is an important task in data analysis and learning. Classification can be performed using supervised or unsupervised methods. From the unsupervised point of view, Formal Concept Analysis (FCA) can be used for such a task in an efficient and well-founded way. From the supervised point of view, emerging patterns rely on pattern mining and can be used to characterize classes of objects w.r.t. a priori labels. In this paper, we present a hybrid classification method which is based both on supervised and unsupervised aspects. This method relies on FCA for building a concept lattice and then detects the concepts whose extents determines classes of objects sharing the same labels. These classes can then be used as reference classes for classifying unknown objects. This hybrid approach has been used in an experiment in chemistry for classifying inhibitors of the c-Met protein which plays an important role in protein interactions and in the development of cancer.

Keywords: Formal Concept Analysis, supervised classification, unsupervised classification, emerging patterns, pattern mining

1 Introduction

In this paper, we present a classification approach based on a combination of knowledge discovery methods which are all interconnected. This approach has to guide two processes, classification and prediction, for analyzing the c-Met receptor protein, a molecule showing an abnormally elevated expression in cancer disease [1]. Activation of this receptor can be inhibited by different biochemical compounds (inhibitors). We collected a group of 100 molecules ("complete set of inhibitors") which are known to be c-Met inhibitors. Inhibitors act on c-Met through a "binding pocket" and an associated "binding mode". We know the binding modes for 30 inhibitors of the dataset (so called "training set"). According to the spatial regions involved in the binding pocket, three main binding modes have been determined: "Type-1", "DFG-out", and "C-Helix-out" (the names are given w.r.t. spatial configuration of proteins). The "Type-1" binding mode is very mixed, probably meaning that it should be divided into more specialized modes. Chemists are working on the definition of a fourth binding mode, close to "Type-1" and termed as "Type-1bis".

 Volume published and copyrighted by its editors.
Local Proceedings in ISBN 978–84–695–5252–0,
Universidad de Málaga (Dept. Matemática Aplicada), Spain.

To ensure the best and adapted inhibition, it is important to know the binding mode of an inhibitor, and this can only be done through chemical experiments, which are long and expensive. Thus, two main questions arise here:

- Is it possible to classify the complete inhibitor set of 100 molecules according to the functionality (based on functional groups) and particular substructures detected in the 30 molecules of the training set?
- Is it possible then to predict the binding mode or "class" of the 70 inhibitors based on the classification of the complete inhibitors set?

For answering the two questions, we introduce a combined classification/prediction process involving supervised and unsupervised classification within the framework of FCA, graph mining and the so-called "Jumping Emerging Patterns" (JEPs). More precisely, we first want to classify a set of molecules (of the training set) according to their structure and their functionality (the functionality determines the behavior of a molecule during reaction and is linked to special substructures called functional groups). For analyzing the structures of the molecules in the training set, we consider molecules as graphs and apply graph mining techniques [2, 3] to extract frequent substructures. Then, these substructures are used as attributes in a formal context where objects are molecules of the training set. This formal context is "augmented" in the sense that each molecule in the training set has a "type" or a "class" according to its binding mode. A concept lattice is built from the formal concept. Moreover, the class information is used for characterizing the concepts whose extents include objects of a single class or binding mode. The intents of these particular concepts are JEPs. Closed JEPs have already been studied in the framework of FCA (see [4–6], where they are called JSM-hypotheses). The set of all JEPs forms a "disjunctive version space" which was related to FCA in [7].

The last step involves a "hierarchical agglomerative clustering" process. Based on the knowledge of JEPs and of functional groups, inhibitors are represented as vectors where components are filled with functional groups and JEPs (55 components where 42 are functional groups and 13 are JEPs). The cosine similarity is used for building a dendrogram which is used for explaining the "proximity" of some inhibitors and for predicting the binding mode of inhibitors for which this information is still unknown.

This classification process which calls for a variety of knowledge discovery methods is totally original and is designed for solving a real-world problem. Here, an original combination of supervised and unsupervised classification works in relation with graph mining and clustering. This shows also the flexibility of the FCA process to be combined with other classification methods for giving actual and substantial results. Experiments are still running but preliminary results have been reached and show that the approach should be continued and improved.

The paper is organized as follows. In Section 2 a motivating example is introduced. Then Section 3 describes the classification flow. Section 4 introduces the main definitions on FCA, JEPs and how they are extracted. Section 5 details

	H	CAD	OH	P	AAE	F	O=
319	x	x			x	x	x
320	x	x		x		x	x
L5G		x	x				x
ZZY	x			x	x	x	

(a) Formal Context

Molecule	Binding Mode
319	DFG-out
320	DFG-out
L5G	Type-1
ZZY	Type-1

(b) Molecule Binding Modes

Table 1: Running Example. In 1a, objects (the rows) are molecules; attributes (the columns) are functional groups. A cross in the cell (i, j) means that the molecule i includes the functional group j as a substructure. In 1b the last column designates the "class" of an object, i.e. the binding mode of the molecule.

the preparation of the molecular data that are processed with FCA. The clustering method and its application are following. The main results are discussed in Section 7 before the conclusion of the paper.

2 Running Example

Formal Concept Analysis (FCA) is briefly introduced hereafter. FCA is based on a formal context which is a triple (G, M, I), where G is a set of objects, M is a set of attributes and $I \subseteq G \times M$ is a relation between G and M [8].

A running example is shown in Table 1. Molecules are objects which are described by substructures, corresponding to attributes. The selection of these particular substructures is discussed later.

Concept	A Set of Molecules (Extent)	A Set of Substructures (Intent)
C_0		H, CAD, OH, P, AAE, F, O=
C_1	ZZY	H, P, AAE, F
C_2	319	H, CAD, P, F, O=
C_3	320	H, CAD, AAE, F, O=
C_4	L5G	CAD, OH, O=
C_5	319, 320	H, CAD, F, O=
C_6	320, ZZY	H, P, F
C_7	319, ZZY	H, AAE, F
C_8	319, 320, L5G	CAD, O=
C_9	319, 320, ZZY	H, F
C_{10}	319, 320, L5G, ZZY	

Table 2: A set of formal concepts w.r.t context on Table 1a.

For every set of molecules A it is possible to find the maximal set of substructures B, included into every molecule from A. This operation is denoted as $(\cdot)'$ with $B = A'$. For example, molecules 319 (BMS_WO/2005/117867_24) and ZZY (UCB_Celltech_azaindole) include the following substructures: H (Halogen),

AAE (Alkyl Aryl Ether) and F (Figure 2a). Dually, for every set of substructures B it is possible to find the maximal set of molecules, including all substructures from B, denoted by $A = B'$. Substructures CAD (Carboxilic Acid Derivative) and O= (Figure 2b) are included into molecules 319 , 320 (molecule BMS_WO/2006-/004636_132) and L5G (Amgen_WO/2008/008539_123). The attribute P stands for substructure Primary Amine while the attribute OH stands for OH-Compound.

The pairs (A, B) –where A is a set of molecules and B is a set of substructures– such that $A' = B$ and $A = B'$ are called "formal concepts". The set A is the extent and the set B is the intent of the concept. The whole set of formal concepts for the running example is given in Table 2.

Formal concepts are partially ordered w.r.t. inclusion of set of objects or of set of attributes: $(A_1, B_1) \leq (A_2, B_2)$ iff $A_1 \subseteq A_2$ or dually $B_2 \subseteq B_1$. This partial ordering gives rise to a concept lattice. Figure 1 shows the concept lattice related to the running example, where reduced notation is used. There are many algorithms for computing formal concepts and the associated concept lattice [9–11].

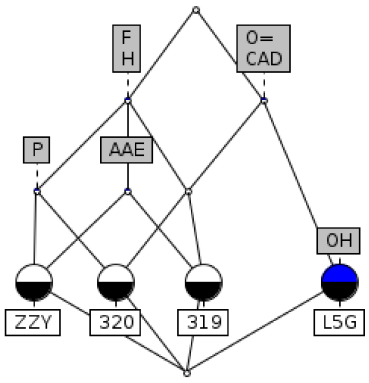
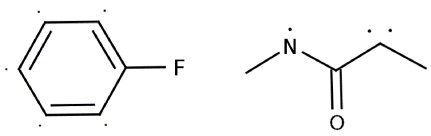


Fig. 1: The FCA-lattice for the context on Table 1.



(a) (b)

Fig. 2: Some substructures for running example in Table 1.

Additional information associated with the molecules is given in Table 1b. The table indicates the binding mode of the molecule with the c-Met protein. This additional column will allow us to process the molecule in a supervised way.

Among concepts in Table 2, it is possible to select concepts whose extent contains only molecules of the same class, e.g. C_1, C_2, C_3, C_4, C_5. The sets of substructures in the intents of these concepts are considered as JEPs and they describe sets of molecules with the same binding mode.

It can be noticed that concept C_5 is more general than concepts C_2 and C_3 since the extent of C_5 includes a wider set of molecules and its intent includes a narrower set of substructures than in extents and intents of C_2 and C_3. As the extent of C_5 only contains molecules of the same type ("DFG-out"), it can be inferred that the substructures in the intent of C_5 characterize this binding mode in a "general and sufficient" way. Accordingly, we are interested in the most general concepts able to describe the binding modes. Here, we obtain concepts C_1, C_4, C_5, and their intents correspond to the most general JEPs.

Since every most general JEP is likely a characteristic of a binding mode, it is worth including these JEPs into molecule descriptions for any clustering or classification purposes. Molecules of the running example can be clustered as shown in Figure 3. This figure should be read as follows: molecules `319` and `320` are close to each other, and are forming a cluster. This cluster is close to molecule `ZZY` and thus molecules `319`, `320`, and `ZZY` are forming a cluster at the next level. Finally, the four molecules are agglomerated into one larger cluster. This clustering process shows the "proximity" of each molecule w.r.t. the binding mode. In this way, clustering can be used to predict the binding mode of an unknown molecule.

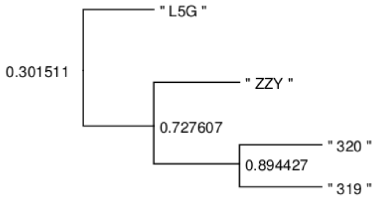


Fig. 3: The clustering result for the context on Table 1a.

3 The Classification Flow

A typical supervised classification task involves a database divided into a training set and a test set. The training set and the test set are sets of objects with their descriptions, where every object of the training set is labeled with a given class. Then a supervised classification method searches for rules in the training set, which can classify objects of the test set.

In our case, the database consists of public and known inhibitors of the c-Met protein. Here we consider 100 molecules, 30 in the training set and 70 in the test set (some molecules are shown in Figure 5). As indicated in the introduction, inhibitors can interact with the c-Met protein w.r.t. three different binding modes, plus one hypothetical binding mode under study [1]. Thus, in this work, four binding modes were used for labeling molecules in the training set. The objective is then to predict the binding modes of the molecules lying in the test set.

Figure 4 depicts the global classification flow. The first step is to choose the way how a molecule should be described. One way is to take into account

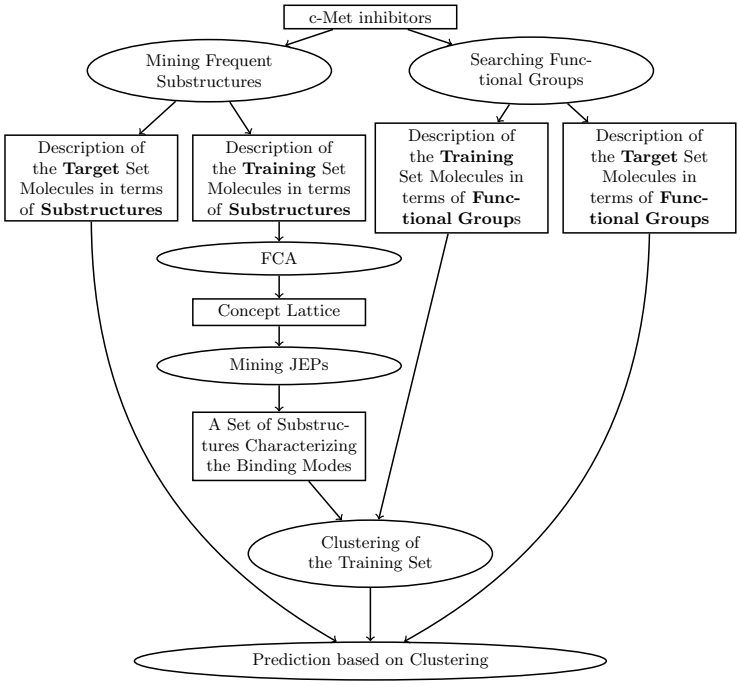


Fig. 4: Diagram of the Classification Flow.

domain knowledge and to consider a molecule as a set of functional groups that are involved into interactions. But some other substructures are also involved into interactions, which are detected as follows:

1. Molecules from a dataset are considered as graphs, where vertexes correspond to atoms and edges to bonds between atoms.
2. A graph miming method is used to find all frequent subgraphs, i.e. subgraphs that belong to a significant part of molecules in the dataset.
3. A formal context is built in the following way:
   - Molecules are considered as objects.
   - Extracted substructures are considered as attributes.
   - A molecule m and a substructure s are related iff the molecule m includes s as a substructure.
4. JEPs (the sets of attributes that characterize only objects of the same class) are extracted from the formal context.

In the supervised classification task, the extracted substructures are used with functional groups to cluster molecules and to predict the binding mode of molecules in the test set.

4 Jumping Emerging Pattens (JEPs)

JEPs were introduced as a means for classification in itemset mining [12, 13], but the underlying idea had appeared and had been studied much earlier, e.g., within the framework of disjunctive version spaces [14, 15] or JSM-hypotheses. Consider an "augmented context", i.e. a context (G, M, I) taken with an additional "class attribute" giving "class information", i.e. the class of each object in G. For a concept (A, B) the set of attributes B is a JEP if every object in A is of the same class. In Table 1, the set of attributes {`F`, `O=`} is a JEP because objects `319` and `320` including these attributes are of the same class "DFG-out".

Since a JEP characterizes a class of objects, it can be used for analyzing this class and for guiding a clustering method. Usually, the set of attributes associated with a single object is trivially a JEP, but there are especially interesting JEPs characterizing a class of objects. The set of JEPs can be partially ordered w.r.t. the subset relation: if there are two JEPs J_1 and J_2 such that J_1 is a subset of J_2, then J_1 is more general, since it describes all the objects described by J_2 and some other objects. For example, the JEP J_1 ={`H`, `CAD`, `F`, `O=`} is more general than the JEP J_2 ={`H`, `CAD`, `P`, `F`, `O=`} since J_2 describes object `320`while J_1 describes objects `320`and `319`.

Relying on the JEP definition, the intent of a formal concept is a JEP if all objects in the concept extent are in the same class. Thus it is possible to compute the set of concepts for a given context and then to extract the JEPs by checking the class of objects in the concept extents. Moreover, the most general JEPs can be selected for further analysis and for clustering.

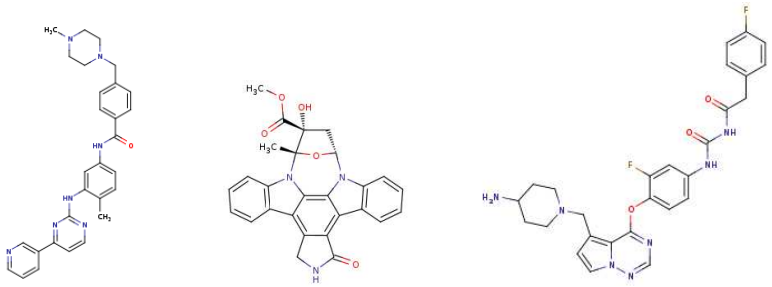
5 Graph Mining

A molecule is a complex structure composed of atoms connected by bonds, that can be considered as a graph. Vertexes of the molecule graph correspond to the atoms of the molecule and are labeled with atom names. The edges of the molecule graph are labeled with types of bonds between the corresponding atoms. For applying FCA and for finding a set of JEPs, a molecular graph can be described as as a set of subgraphs. Then, a formal context can be built with G as a set of molecules, M as a set of subgraphs or substructures and I the relation meaning that a molecule g has a substructure m. The problem now is to find "valid" and "interesting" substructures.

One way to select valid and interesting substructures is to search for frequent subgraphs –that often appear in molecular graphs– using graph mining. For a set of graphs G and a frequency threshold F_{min}, a graph s is frequent iff s is a subgraph of at least F_{min} graphs from G, i.e. $|\{g \in G \mid s \subseteq g\}| \geq F_{min}$.

For example, considering the set of molecular graphs G in Figure 5 and $F_{min} = 3$, the subgraphs "N-H" and "O=C" are frequent as they occur in all molecular graphs while subgraph "C-OH" only occurring in graph (b) (Figure 5b) and subgraph "F-C" only occurring in graph (c) (Figure 5c) are not frequent.

For discovering frequent subgraphs, different graph mining algorithms may be applied [2, 3]. Here we used `gSpan` and set $F_{min} = 10$ for the dataset of 100



(a) Imatinib or Gleevec® (b) K-252a (c) CKK

Fig. 5: Examples of molecules from database.

molecular graphs. This frequency threshold is sufficiently low to have a set of specific subgraphs characterizing every molecule, and it is sufficiently high to obtain feasible processing time.

The set of mined subgraphs can be divided into groups, where a group consist of a set of subgraphs appearing in the same set of molecular graphs. Thus, the group forms an equivalence class and can be represented by only one subgraph. Furthermore, the largest subgraphs preserve the sufficient information on substructures related to binding modes. In the present experiment, around 10^6 frequent subgraphs were extracted, then divided into 10^4 groups.

It can be noticed that if there are two frequent subgraphs g_1 and g_2 such that $g_1 \subseteq g_2$ then every closed JEP containing the subgraph g_2 contains the subgraph g_1. Thus, if a JEP contains g_2, there is no need to consider g_1.

6 Hierarchical Agglomerative Clustering (HAC)

Here we describe a hierarchical agglomerative clustering process (see [16]) based on the extracted JEPs and background knowledge on functional groups. Molecules are described by vectors having 55 components, including 42 functional groups[3] and 13 JEPs. The 13 JEPs are selected as the most representative for the molecules in the training set. Each attribute of the vector therefore corresponds either to a chemical functional group or to a substructures of a JEP with value set to 1 when this chemical function/substructure is present and *null* otherwise. The choice of a proper similarity is crucial for ensuring the quality of the clustering. Here, the cosine similarity was chosen according to the results of several specialized studies [17, 18]. If m_1 and m_2 are the description vectors of two molecules, then $((\boldsymbol{m_1}, \boldsymbol{m_2})$ denotes the scalar product of two vectors):

[3] The functional groups were extracted thanks to the specialized algorithm "Checkmol" http://merian.pch.univie.ac.at/ nhaider/cheminf/cmmm.html.

$$sim_{cos}(\boldsymbol{m_1}, \boldsymbol{m_2}) = \frac{(\boldsymbol{m_1}, \boldsymbol{m_2})}{|\boldsymbol{m_1}| \cdot |\boldsymbol{m_2}|} \tag{1}$$

The "centroid" of a cluster of molecules C, denoted by $\boldsymbol{centr}(C)$, is calculated as follows:

$$\boldsymbol{centr}(C) = \frac{1}{|C|} \sum_{m_i \in C} \boldsymbol{m_i} \tag{2}$$

Similarity between two clusters or between a molecule and a cluster is calculated with the same formula (1) by substituting the cluster C with its centroid $\boldsymbol{centr}(C)$.

The HAC clustering is a bottom-up process working as follows. For every molecule a unique cluster is created. Actually, all these clusters will be progressively merged until only one unique cluster remains. Considering at some step the set of remaining clusters $\mathfrak{C} = \{C_1, C_2, .., C_k\}$, a new cluster C_{k+1} is created by merging the two clusters C_i and C_j maximizing the similarity measure between them. The new cluster is added to the set of clusters while C_i and C_j are deleted from $\mathfrak{C}$. Finally, the process stops when only one cluster remains, $|\mathfrak{C}| = 1$.

$$(C_i, C_j) = \underset{C_i, C_j \in \mathfrak{C}, C_i \neq C_j}{argmax} sim_{cos}(\boldsymbol{centr}(C_i), \boldsymbol{centr}(C_j)) \tag{3}$$

$$C_{k+1} := C_i \cup C_j \tag{4}$$

$$\mathfrak{C} := \mathfrak{C} \cup \{C_{k+1}\} \setminus \{C_i, C_j\} \tag{5}$$

The result of HAC is shown on a dendrogram (see Figures 3 and 6). Each "vertex" of the dendrogram corresponds to a merging step of the algorithm. The number attached to the vertex represents the similarity between the two clusters at the lower level. The correlation between chemical similarities and binding modes is discussed below.

7 Results and Discussion

After applying graph mining on the set of molecules, a formal context including 30 objects (molecules) and 10^4 attributes (substructures) was built. The cardinalitiy of the sets of most general JEPs for the different binding modes are distributed as follows:

- 35 JEPs for Type-1 binding mode;
- 1 JEP for DFG-out binding mode;
- 1 JEP for C-Helix-out binding mode;
- 3 JEPs for Type-1bis binding mode.

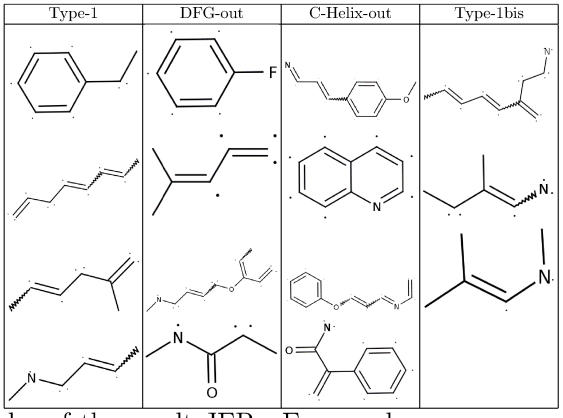


Table 3: Examples of the result JEPs. Every column corresponds to one JEP. Only some of structures for every JEP were exemplified. According to the dataset, all the molecules including all the substructures of the second (for example) column are of DFG-out binding mode. These sets of substructures belongs to disjoint sets of molecules.

Examples of extracted JEPs for different binding modes are shown in Table 3. Substructures associated with the most general JEPs were used in the description of a molecule for the clustering. A molecule was described by a set of functional groups and by the set of JEPs extracted by the mining process. The resulting dendrogram is shown in Figure 6. Two small clusters (0.485071 and 649934) are covering Type-1 molecules, while one small cluster (0.673565) is covering DFG-out molecules and a quite large cluster (0.681201) is covering DFG-out molecules as well as C-Helix-out molecules and one Type-1 molecule. The C-Helix-out molecules may appear in that cluster since they are quite similar while this Type-1 molecule share some chemical properties with the DFG-out molecules. It should be noticed that the dendrogram shows a better cohesion for molecules of the same binding mode and a better separation between molecules of different binding modes, than the dendrogram built from molecules only described by functional groups, and this is mainly due to JEPs substructures,

An extended dendrogram will be tested for the set of 100 molecules to check whether the class of some unknown molecule can be determined with respect to its "proximity" to other molecules in the dendrogram.

8 Conclusion

In this paper, we have shown how to classify a set of molecules with respect to their structure. Actually, we combined two classification aspects: supervised and unsupervised classifications. First, molecules are represented by molecular graphs and graph mining is applied to these graphs for extracting interesting substructures. Then, FCA is applied on an "augmented" context where there

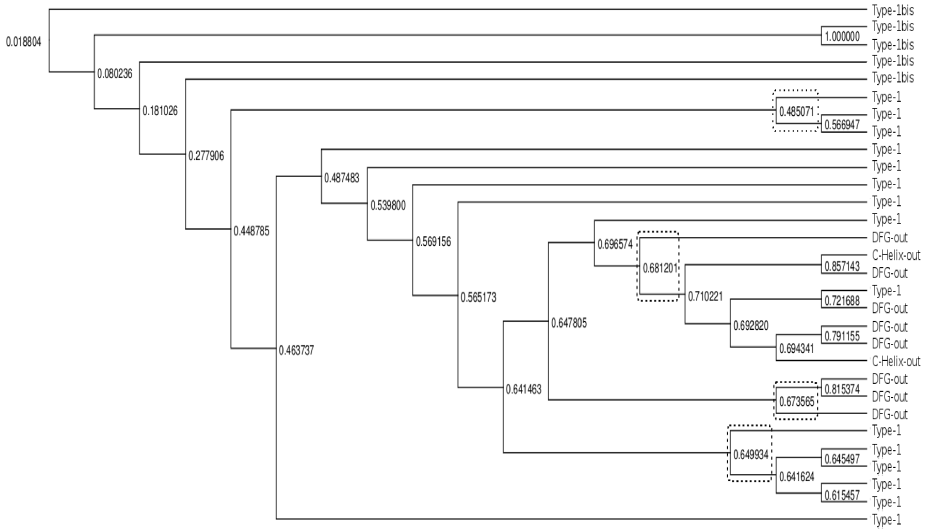


Fig. 6: The clustering result for 30 known molecules, described by functional groups and JEPs.

is a "class" information for objects. This allows one to extract JEPs. Then, JEPs and functional groups are used to cluster the molecules and to obtain a dendrogram showing the proximity between molecules.

In future work, we are going to use the dendrogram for prediction with the supervision of chemists. There are several ways to do that. One way to classify an unknown molecule is to find the closest cluster of molecules sharing the same label. Another way is to cluster all the molecules (complete set of inhibitors) and then to predict the class of unknown molecules, using information in the common class of molecules in the same cluster.

Another direction of the future work is to study and apply a knowledge-based evaluation function in graph mining, for selecting the relevant subgraphs from the chemical point of view instead of frequent ones. Finally, following [19], we are planning to combine a similarity-based approach with FCA, to avoid the clustering step and to build a lattice including all interesting information.

References

1. Asses, Y., Leroux, V., Tairi-Kellou, S., Dono, R., Maina, F., Maigret, B.: Analysis of c-met kinase domain complexes: A new specific catalytic site receptor model for defining binding modes of ATP-Competitive ligands. Chemical Biology & Drug Design **74**(6) (2009) 560–570
2. Yan, X., Han, J.: gSpan: graph-based substructure pattern mining. In: Proceedings of ICDM'02. (2002) 721 – 724
3. Nijssen, S., Kok, J.: The gaston tool for frequent subgraph mining. Electronic Notes in Theoretical Computer Science **127** (March 2005) 77–87

4. Ganter, B., Kuznetsov, S.: Formalizing hypotheses with concepts. In Ganter, B., Mineau, G., eds.: Conceptual Structures: Logical, Linguistic, and Computational Issues. Volume 1867 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2000) 342–356
5. Blinova, V.G., Dobrynin, D.A., Finn, V.K., Kuznetsov, S.O., Pankratova, E.S.: Toxicology analysis by means of the JSM-method. Bioinformatics **19**(10) (2003) 1201–1207
6. Kuznetsov, S.O., Samokhin, M.V.: Learning closed sets of labeled graphs for chemical applications. In: Proceedings of ILP. (2005) 190–208
7. Ganter, B., Kuznetsov, S.: Hypotheses and version spaces. In Ganter, B., de Moor, A., Lex, W., eds.: Conceptual Structures for Knowledge Creation and Communication. Volume 2746 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2003) 83–95
8. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. 1st edn. Springer-Verlag New York, Inc., Secaucus, NJ, USA (1997)
9. Ganter, B.: Two basic algorithms in concept analysis. In Kwuida, L., Sertkaya, B., eds.: Formal Concept Analysis. Volume 5986 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (1984) 312–340
10. Kuznetsov, S.O.: A fast algorithm for computing all intersections of objects in a finite semi-lattice. Automatic documentation and Mathematical linguistics **27**(5) (1993) 11–21
11. Merwe, D., Obiedkov, S., Kourie, D.: AddIntent: a new incremental algorithm for constructing concept lattices. In Goos, G., Hartmanis, J., Leeuwen, J., Eklund, P., eds.: Concept Lattices. Volume 2961. Springer Berlin / Heidelberg (2004) 372–385
12. Dong, G., Li, J.: Efficient mining of emerging patterns: discovering trends and differences. In: Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '99, New York, ACM (1999) 43–52
13. Poezevara, G., Cuissart, B., Crémilleux, B.: Extracting and summarizing the frequent emerging graph patterns from a dataset of graphs. Journal of Intelligent Information Systems **37** (July 2011) 333–353
14. Sebag, M.: Delaying the choice of bias: A disjunctive version space approach. In: Proceedings of the 13 th International Conference on Machine Learning, Morgan Kaufmann (1996) 444–452
15. Nikolaev, N.I., Smirnov, E.N.: Stochastically guided disjunctive version space learning. In: Proceedings of the 12th European Conference on Artificial Intelligence, John Wiley & Sons, Ltd. (1996)
16. Murtagh, F.: A survey of recent advances in hierarchical clustering algorithms. The Computer Journal **26**(4) (November 1983) 354–359
17. Qian, G., Sural, S., Gu, Y., Pramanik, S.: Similarity between euclidean and cosine angle distance for nearest neighbor queries. In: Proceedings of 2004 ACM Symposium on Applied Computing, ACM Press (2004) 1232–1237
18. Yamagishi, M., Martins, N., Neshich, G., Cai, W., Shao, X., Beautrait, A., Maigret, B.: A fast surface-matching procedure for protein–ligand docking. Journal of Molecular Modeling **12**(6) (2006) 965–972
19. Kaytoue, M., Assaghir, Z., Napoli, A., Kuznetsov, S.O.: Embedding tolerance relations in formal concept analysis: an application in information fusion. In: Proceedings of the 19th ACM international conference on Information and knowledge management. CIKM '10, New York, NY, USA, ACM (2010) 1689–1692