From bags of graphs to stereo subgraphs in order to predict molecule's properties

Pierre-Anthony Grenier, Luc Brun, D. Villemin

From bags to graphs of stereo subgraphs in order to predict molecule's properties

Pierre-Anthony Grenier[†], Luc Brun[†], and Didier Villemin[‡]

†GREYC UMR CNRS 6072, ‡LCMT UMR CNRS 6507,
Caen, France
{pierre-anthony.grenier,luc.brun,didier.villemin}@ensicaen.fr,

Abstract. Quantitative Structure Activity and Property Relationships (QSAR and QSPR), aim to predict properties of molecules thanks to computational techniques. In these fields, graphs provide a natural encoding of molecules. However some molecules may have a same graph but differ by the three dimensional orientation of their atoms in space. These molecules, called stereoisomers, may have different properties which cannot be correctly predicted using usual graph encodings. In a previous paper we proposed to encode the stereoisomerism property of each atom by a local subgraph. A kernel between bags of such subgraphs then provides a similarity measure incorporating stereoisomerism properties. However, such an approach does not take into account potential interactions between these subgrahs. We thus propose in this paper, a method to take these interactions into account hence providing a global point of view on molecules's stereoisomerism properties.

Keywords: Graph kernel, Chemoinformatics, Stereoisomerism.

1 Introduction

QSAR and QSPR methods are based on a basic principle which states that: "two similar molecules should have similar properties". An usual way to encode molecules is to use their molecular graphs. A molecular graph is a simple graph $G = (V, E, \mu, \nu)$, where each node $v \in V$ encodes an atom, each edge $e \in E$ encodes a bond between two atoms and the labeling functions μ and ν associate to each vertex and each edge a label encoding respectively the nature of the atom (carbon, oxygen,...) and the type of the bond (single, double, triple or aromatic).

However, molecular graphs have a limitation: they do not encode the spatial configuration of atoms. Indeed, some molecules, called stereoisomers, are associated to a same molecular graph but differ by the relative positioning of their atoms. We can imagine for example, a carbon atom, with four neighbors, each of them located on a summit of a tetrahedron. If we permute two of the atoms, we obtain a different spatial configuration (Figure 1a). An atom is called a stereocenter if a permutation of two atoms belonging to its neighborhood produces a different stereoisomer. Two connected atoms also define a stereocenter if a

(a) Two different spatial configurations of the neighbors of a carbon

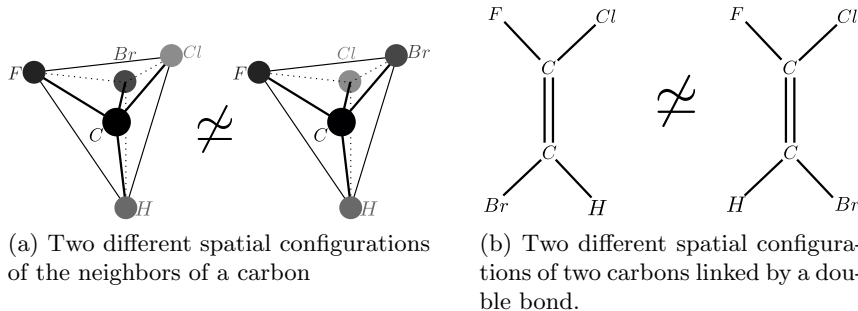(b) Two different spatial configurations of two carbons linked by a double bond.

Fig. 1: Two types of stereocenters.

permutation of the positions of two atoms belonging to the union of their neighborhoods produces a different stereoisomer (Figure 1b). According to chemical experts [8], within molecules currently used in chemistry, 98% of stereocenters correspond either to carbons with four neighbors, called asymmetric carbons (Figure 1a) or to couples of two carbons adjacent through a double bond (Figure 1b). We thus restrict the present paper to such cases.

Graph kernels [9–11, 3], allow us to combine a graph encoding of molecules with usual machine learning methods. Up to now, only few methods have attempted to incorporate stereoisomerism within the graph kernel framework. Brown et al. [1] have proposed to incorporate this information through an extension of the tree-pattern kernel [10]. In this last method, similarity between molecules is deduced from the number of common tree-patterns between two molecules. When several stereocenters are close to each other, one pattern may implicitly encode a walk which connect them. However the size of patterns being limited, in some cases the influence of a permutation around stereocenters is not detected.

Intuitively, stereoisomerism property is related to the fact that permuting two neighbors of a stereocenter produces a different spatial configuration. If those two neighbors have a same label, the influence of the permutation should be searched beyond the direct neighborhood of this stereocenter. Based on this ascertainment, we have proposed in [7] to characterized a stereocenter by a subgraph, big enough to highlight the influence of each permutation of the neighbors of this stereocenter but sufficiently small to provide a local characterization of it. We then proposed a kernel based on those subgraphs.

One drawback of our previous approach is that each subgraph, and thus each stereocenter, is considered independently. We thus present in this paper a method based on [7], which explicitly encode all minimal stereo subgraphs associated to stereocenters together with several types of interactions between these subgraphs.

In Section 2 we remind the two main points of [7], the encoding of molecules by ordered graphs, and the construction of minimal stereo subgraphs which characterize stereocenters. Then in Section 3 we present new graph models taking

into account relationships between minimal stereo subgraphs. Results obtained with those graphs are provided in Section 4.

2 Ordered graphs and minimal stereo subgraphs

2.1 Encoding of molecules by ordered graphs

The spatial configuration of the neighbors of each atom may be encoded through an ordering of its neighborhood [7]. In order to encode this information, we introduce the notion of ordered graph. An ordered graph $G = (V, E, \mu, \nu, ord)$ is a molecular graph $G_m = (V, E, \mu, \nu)$ together with a function $ord : V \to V^*$ which maps each vertex to an ordered list of its neighbors. Two ordered graphs G and G' are isomorphic ($G \underset{o}{\simeq} G'$) if there exists an isomorphism f between their respective molecular graphs G_m and G'_m such that $ord'(f(v)) = (f(v_1) \ldots f(v_n))$ with $ord(v) = (v_1 \ldots v_n)$ (where $N(v) = \{v_1, \ldots, v_n\}$ denotes the neighborhood of v). In this case f is called an ordered isomorphism between G and G'.

However, different ordered graphs may encode a same molecule. We thus have to define an equivalence relationship between ordered graphs, such that two ordered graphs are equivalent if they represent a same molecular configuration.

To do so, we introduce the notion of re-ordering function σ, which associates to each vertex $v \in V$ of degree n a permutation $\sigma(v)$ on $\{1, \ldots, n\}$, which allows to re-order its neighborhood. The graph with re-ordered neighborhoods $\sigma(G)$ is obtained by mapping for each vertex v its order $ord(v) = v_1 \ldots . v_n$ onto the sequence $v_{\sigma(v)(1)} \ldots . v_{\sigma(v)(n)}$ where $\sigma(v)$ is the permutation applied on v.

The set of re-ordering functions, transforming an ordered graph into another one representing the same configuration is called a valid family of re-ordering functions Σ [4]. We say that it exists an equivalent ordered isomorphism f between G and G' according to Σ if it exists $\sigma \in \Sigma$ such that f is an ordered isomorphism between $\sigma(G)$ and G' ($\sigma(G) \underset{o}{\simeq} G'$). The equivalent order relationship defines an equivalence relationship [4] and two different stereoisomers are encoded by non equivalent ordered graphs. We denote by $\mathrm{IsomEqOrd}(G, G')$ the set of equivalent ordered isomorphism between G and G'.

Carbons with four neighbors, and double bonds between carbons, are not necessarily stereocenters. If they are not stereocenters, any permutation in their neighbourhood would lead to an equivalent ordered graph. We thus define for an ordred graph $G = (\widehat{G} = (V, E, \mu, \nu), ord)$ and one of its vertex $v \in V$ a set of ordrered isomorphism $\mathcal{F}_G^v$:

$$\mathcal{F}_G^v = \bigcup_{\substack{(i,j) \in \{1, \ldots, |N(v)|\}^2 \\ i \neq j}} \{f \mid f \in \mathrm{IsomEqOrd}(G, \tau_{i,j}^v(G)) \text{ with } f(v) = v\}$$

where $\tau_{i,j}^v$ is a re-ordering function equals to the identity on all vertices except v for which it permutes the vertices of index i and j in $ord(v)$.

We then define a stereo vertex as a vertex for which any permutation of two of its neighbors produces a non-equivalent ordered graph:

Definition 1 (Stereo vertex). Let $G = (V, E, \mu, \nu, ord)$ be an ordered graph. A vertex $v \in V$ is called a stereo vertex iff $\mathcal{F}_G^v = \varnothing$.

Two carbons linked by a double bond form a stereocenter and we have proved in [6] that if a carbon of a double bond is a stereo vertex then the other one is also a stereo vertex. Therefore we denote by $kernel(s)$ the set of stereo vertices corresponding to a stereocenter ($kernel(s) = \{s\}$ if s is an asymmetric carbon and $kernel(s) = \{s, u\}$ if s is a carbon of a double bond, where u is the other carbon of the double bond). We further denote by $StereoStar(s)$ the set composed of a stereocenter and its neighbourhood: $StereoStar(s) = kernel(s) \cup N(kernel(s))$.

2.2   Minimal stereo subgraphs

Definition 1 is based on the whole graph G to test if a vertex v is a stereo vertex. However, given a stereo vertex s, one can observe that on some configurations, the removal of some vertices far from s should not change its stereo property. In order to obtain a more local characterization of a stereo vertex, we should thus determine a vertex induced subgraph H of G, including s, large enough to characterize the stereo property of s, but sufficiently small to encode only the relevant information characterizing the stereo property of s. Such a subgraph is called a minimal stereo subgraph of s.

We now present a constructive definition of a minimal stereo subgraph of a stereo vertex. Let s denote a stereo vertex and let H_s be a subgraph of G containing $kernel(s)$. We say that the stereo property of s is not captured by H_s if (Definition 1):

$$\mathcal{F}_{H_s}^s \neq \varnothing \tag{1}$$

To define a minimal stereo subgraph of s, we consider a finite sequence $(H_s^k)_{k=1}^n$ of vertex induced subgraphs of G. The first element of this sequence H_s^1 is the smallest vertex induced subgraph for which we can test (1): $V(H_s^1) = StereoStar(s)$.

If the current vertex induced subgraph H_s^k does not capture the stereo property of s, we know by (1), that it exists some isomorphisms $f \in \mathcal{F}_{H_s^k}^s$. We denote by $\mathcal{E}_f^k$ the set of vertices of H_s^k inducing the isomorphism f in H_s^k:

$$\mathcal{E}_f^k = \{v \in V(H_s^k) \mid \exists p = (v_0, \ldots, v_q) \in H_s^k \text{ with } v_0 \in kernel(s) \text{ and } v_q = v$$
$$\text{s.t. } f(v_1) \neq v_1\} \tag{2}$$

In [6], we show that for any f in $\mathcal{F}_{H_s^k}^s$, $\mathcal{E}_f^k$ is not empty. A vertex v belongs to $\mathcal{E}_f^k$ if neither its label nor its neighborhood in H_s^k allow to differentiate it from $f(v)$. The basic idea of our algorithm consists in enforcing constraints on each $v \in \mathcal{E}_f^k$ at iteration $k + 1$ by adding to H_s^k the neighborhood of v in G. The set of vertices of the vertex induced subgraph H_s^{k+1} is thus defined by:

$$V(H_s^{k+1}) = V(H_s^k) \cup \bigcup_{f \in \mathcal{F}_{H_s^k}^s} N(\mathcal{E}_f^k) \tag{3}$$
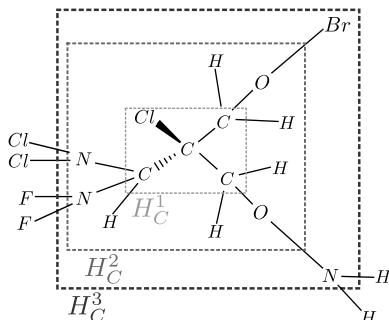
Fig. 2: An asymmetric carbon and its associated sequence $(H_C^k)_{k=1}^3$

where $N(\mathcal{E}_f^k)$ denote the neighborhood of $\mathcal{E}_f^k$.

The algorithm stops when the set $f \in \mathcal{F}_{H_s^k}^s$ becomes empty. We proved in [6] that the subgraph obtained by this algorithm captures the stereo property of s. Figure 2 illustrates our algorithm. Remarks that the computation of the minimal stereo subgraph requires the computation of graph isomorphisms and is thus nearly NP-complete, however note that stereo subgraphs correspond to a local characteristic of vertex and have consequently a limited size [7].

Thus for each stereo vertex we can construct its minimal stereo subgraph to characterize it. We consider two stereo vertices as similar if they have a same minimal stereo subgraph, and to test it efficiently, we transform our minimal stereo subgraphs S into codes c_S thanks to the method described in [13].

3   Graph of interactions

In the previous section we have defined a way to encode molecules and construct an oriented subgraph which characterizes a stereocenter. We associate to an

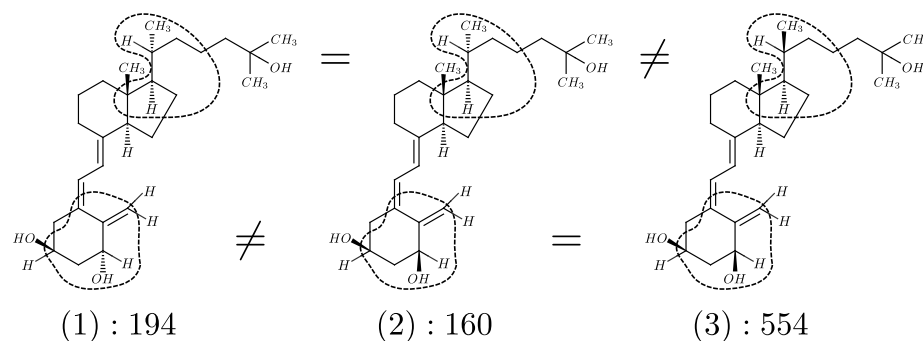$(1):194$ $\qquad$ $(2):160$ $\qquad$ $(3):554$

Fig. 3: 3 molecules with the value of their biological activities. Minimal stereo subgraphs which differs between them are surrounded by dotted lines.

ordered graph G its bag of minimal stereo subgraph $\mathcal{H}(G)$. In [7], we have proposed a kernel between these bags which encodes a similarity between ordered graphs. However, using this kernel, each minimal stereo subgraph is considered independently.

Figure 3 shows an example of three molecules of a dataset used in Section 4. In this figure, (2) have only one minimal stereo subgraph different from (1) and (3). Thus by considering the notion of distance associated to the kernel of [7], (1) and (3) are equidistant from (2). However the biological activity of (2) is closer from the biological activity of (1) than from the one of (3). As the different minimal stereo subgraph between (2) and (3) is close from other minimal stereo subgraphs, taking into account interactions between minimal stereo subgraphs allows to obtain a smaller distance between (1) and (2) than between (3) and (2), which may help to obtain a better prediction of the property.

Unfortunately, the amount of interactions between two stereo subgraphs which may influence a molecular property is yet unknown both in chemoinformatics and chemistry fields. Hence we propose to define different functions of interactions, encoding different degrees of information about the interactions between stereo vertices.

Functions of interactions are defined according to a sequence of conditions $(c_1, \ldots, c_n)$. These conditions are increasingly constraining:

$$\forall i \in \{1, \ldots, n-1\}\, c_{i+1} \Rightarrow c_i$$

Let H_1 and H_2 be two minimal stereo subgraphs, such that s_1 is the stereo vertex of H_1 and s_2 is the stereo vertex of H_2. We propose the following set of conditions:

$$c_1 : H_1 \bigcap H_2 \neq \emptyset \qquad c_2 : kernel(s_1) \subset H_2$$
$$c_3 : StereoStar(s_1) \subset H_2 \quad c_4 : H_1 \subset H_2$$

We consider in this paper three functions of interactions F_i. Each function F_i is designed in order to be more restrictive than F_j (with $j < i$). To do so each F_i is defined by only using conditions c_j with j in $\{i, \ldots, 4\} \cup \{0\}$, where c_0 is defined as $\neg c_i$. The value $F_i(H_1, H_2)$ is obtained by taking the maximal index j of conditions c_j which represents the strongest interaction between H_1 and H_2:

$$F_i(H_1, H_2) = \max\{j \in \{i, \ldots, 4\} \cup \{0\} \mid c_j\}$$

Note that $(F_i)_{i \in \{1,2,3\}}$ are non symmetric functions.

We define thanks to those functions, three graphs of interactions G_i where each vertex $v \in V_i$ represents a minimal stereo subgraph and each edge encodes an interaction between two minimal stereo subgraphs deduced from F_i :

Definition 2 (Graph of interactions). A graph of interactions $G_i = (V_i, E_i, \mu_i, \nu_i)$ is a graph built from an ordered graph $G = (G_m = (V, E, \mu, \nu), ord)$ and a function of interaction F_i such that :

- $\forall u \in V_i, \exists! H(u) \in \mathcal{H}(G)$.
- $\forall u \in V_i, \mu_i(u) = c_{H(u)}$, where c_H is the code defined in [13] (Section 2).
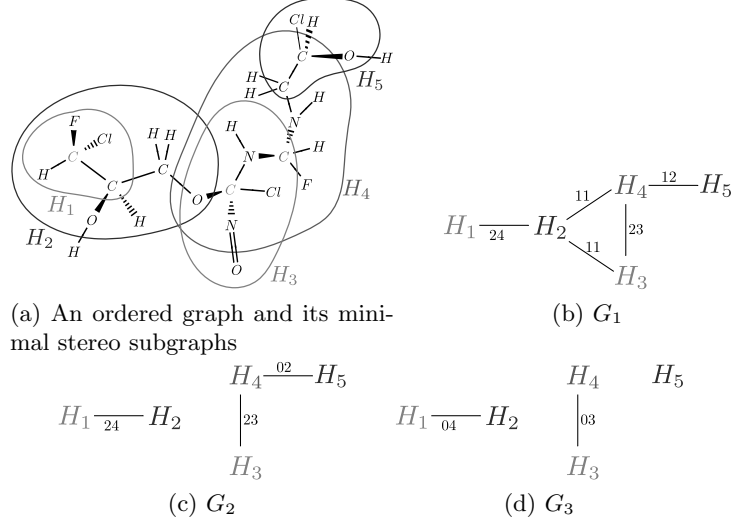
(a) An ordered graph and its minimal stereo subgraphs

(b) G_1

(c) G_2

(d) G_3

Fig. 4: One ordered graph and its different graphs of interactions G_i, obtained using F_i with $i \in \{1, 2, 3\}$

- $\exists (u_1, u_2) \in E_i \iff F_i(H(u_1), H(u_2)) \neq 0$ or $F_i(H(u_2), H(u_1)) \neq 0$.
- $\forall e = (u_1, u_2) \in E_i, \; \nu_i(e) = \min(F_i(H(u_1), H(u_2)), F_i(H(u_2), H(u_1))) \odot \max(F_i(H(u_1), H(u_2)), F_i(H(u_2), H(u_1)))$.

where $\odot$ denotes the concatenation.

We can check if a vertex is in a minimal stereo subgraph in constant time. Thus, the complexity to check each conditions considering a minimal subgraph H_1 with H_2 is $\mathcal{O}(max(|H_1|, |H_2|))$ for c_1, $\mathcal{O}(|kernel(s_1)|)$ for c_2, $\mathcal{O}(|StereoStar(s_1)|)$ for c_3 and $\mathcal{O}(|H_1|)$ for c_4. Thus the complexity in the worst case for computing the graphs of interactions is $\mathcal{O}(|\mathcal{H}(G)|^2 \max_{H \in \mathcal{H}(G)} |H|)$. In practice this value is small (for the vitamin dataset presented in Section 4, we have at most $|\mathcal{H}(G)| = 9$ and $\max_{H \in \mathcal{H}(G)} |H| = 24$).

Figure 4 shows the graphs of interactions obtained from an ordered graph using the different functions of interactions. The graph G_1 is built by taking all conditions. However we may suppose that the weaker interactions c_1 may not be relevant. Indeed, an intersection between two minimal stereo subgraphs may not be a sufficiently relevant information. Thus the graph G_2 is designed by considering that two stereo vertices are related if we have at least $kernel(s_1) \subset H_2$ or $kernel(s_2) \subset H_1$. Moreover, a vertex s_1 is a stereo vertex because of the relative positioning of its neighbour. So we may suppose that, if a stereo vertex is present in a stereo subgraph ($kernel(s_1) \subset H_2$), but not its neighbourhood ($StereoStar(s_1) \not\subset H_2$), the stereo vertex may have a similar influence in H_2 than a non-stereo vertex. Thus G_3 is built by considering that two stereo vertices are related if we have at least $StereoStar(s_1) \subset H_2$ or $StereoStar(s_2) \subset H_1$.

Table 1: Average values of numbers of vertices ($\overline{|V|}$), number of edges ($\overline{|E|}$), number of labels($\overline{|\mathcal{L}_V|}$,$\overline{|\mathcal{L}_E|}$), and mean degree ($\overline{d}$) of graph of interactions.

(a) ACE dataset

	$\overline{|V|}$	$\overline{|E|}$	$\overline{|\mathcal{L}_V|}$	$\overline{|\mathcal{L}_E|}$	$\overline{d}$
Graph 1	5	7	4.5	3	2.8
Graph 2	5	2	4.5	2	0.8
Graph 3	5	1	4.5	1	0.4

(b) Vitamin dataset

	$\overline{|V|}$	$\overline{|E|}$	$\overline{|\mathcal{L}_V|}$	$\overline{|\mathcal{L}_E|}$	$\overline{d}$
Graph 1	8.55	17.4	8.38	5.71	4.07
Graph 2	8.55	11.3	8.38	4.71	2.62
Graph 3	8.55	6.14	8.38	2.71	1.43

Table 2: Classification of the ACE inhibitory activity of perindopirilates stereoisomers

Method	Accuracy				
Brown [1]	*96.875*				
Stereo Kernel [7]	87.5				
Stereo + Extended subgraphs [5]	*96.875*				
Graph of interactions with	[3]	[9]	[10]	[11]	[3] with MKL
Graph of interactions 1	93.75	84.375	93.75	93.75	**100**
Graph of interactions 2	93.75	62.5	62.5	62.5	87.5
Graph of interactions 3	84.375	62.5	62.5	62.5	90.625

As graphs of interaction are graphs without order, we may apply any graph kernel (e.g. [9–11, 3]) to measure their similarities. Note that for the treelet kernel [3], treelets of size 1 are vertices of graphs of interactions, and thus encode the same notion of similarity than the bags of stereo subgraphs [7].

4   Experiments

We have tested our method on two datasets. For both of them we use the same protocol: a nested cross-validation which select parameters and estimate the performance. The outer cross-validation is a leave-one-out procedure, used to compute an error of prediction for each molecule of the dataset. For each fold, we use another leave-one-out procedure on the remaining molecules, to compute a validation error. We use standard SVM methods for classification and regression of molecules. Basic statisistics about the graphs of interactions $G_i = (V_i, E_i, \mu_i, \nu_i)$ deduced from each dataset is displayed in Table 1.

Our first experiment is based on a dataset composed of all the stereoisomers of the perindoprilate [2]. As this molecule has 5 stereocenters, the dataset is composed of $2^5 = 32$ molecules. In this dataset, we try to predict if a molecule inhibit the angiotensin-converting enzyme (ACE).

In this dataset two stereocenters have a same minimal stereo subgraph, but different surrounding. The stereo kernel [7] and one of the graph of interactions (G_3), can not differentiate those two stereocenters, which have different influence

Table 3: Prediction of the biological activity of synthetic vitamin D derivatives.

Method	RMSE			
1 - Tree patterns Kernel [10]	0.251			
2 - Treelet Kernel [3]	0.271			
3 - Brown [1]	0.184			
4 - Stereo Kernel [7]	0.194			
5 - Stereo + Extended subgraphs [5]	0.180			
Graph of interactions with	[3]	[9]	[10]	[11]
6 - Graph of interactions 1	0.177	0.184	0.185	0.201
7 - Graph of interactions 2	0.169	0.189	*0.162*	0.166
8 - Graph of interactions 3	0.172	0.172	**0.161**	*0.162*

on the property, this explains why other method ([1, 5] and the two other graphs of interactions) obtain a better accuracy. However, for our graphs of interactions, treelet of size one have a negative effect on the classification, this explains why we do not obtain better results than [1, 5]. By using a multiple kernel learning algorithm [12], we can learn a weight for each treelet, that allow us to discard treelet of size 1 and to obtain the best results with the first graph of interactions. The second graph of interactions have very few edges and a low degree (Table 1a). This last point explains why some kernel [9–11] and the treelet kernel with multiple kernel learning obtains poor results on this graph.

The second dataset is a dataset of synthetic vitamin D derivatives, used in [1]. This dataset is composed of 69 molecules, with an average of 8.55 stereocenters per molecule. This dataset is associated to a regression problem, which consists in predicting the biological activity of each molecule.

Methods which do not encode stereoisomerism information [10, 3] obtain poor results as we can see in Table 3 (lines 1-2). The adaptation of the tree pattern kernel to stereoisomerism [1] and our previous kernels [7, 5] (lines 3-5) improves the results over the two previous methods hence showing the insight of adding stereoisomerism information. Taking into account relationships between minimal stereo subgraphs (lines 6-8) allows us to obtain better results than our previous method [5]. Unlike the previous dataset, the graphs of interactions G_2 and G_3 have higher degree (≈ 2) and thus obtains good results. Graph G_1 have a high degree (4) and a high number of different labels on vertices, which induces a lot of unique patterns in each graph. This last point decreases the number of patterns common to two graphs and explains why on this dataset G_1 does not obtain results as good as G_2 and G_3.

In conclusion, the treelet kernel applied on graphs of interactions seems to obtain equivalent or better results than alternative kernel methods. Moreover, given a data set of molecules related to stereoisomerism, our experiments show that the choice of a particular graph of interactions should be based on the mean degree of vertice of these graphs. A mean degree of 2 seems to correspond to a good compromise between a very low degree corresponding to a node set without graph structure and a too high degree which also hidden the graph structure.

10   Pierre-Anthony Grenier[†], Luc Brun[†], and Didier Villemin[‡]

5   Conclusion

In this paper we have proposed an extension of our previous method [7] based on a new graph model, where each node represents a stereo subgraph and each edge encodes an interaction between two stereo subgraphs. This graph allows us to take into account interactions between stereo subgraphs. The relevance of this approach is demonstrated through experiments on two datasets.

Acknowledgements

References

1. J. Brown, T. Urata, T. Tamura, M. A. Arai, T. Kawabata, and T. Akutsu. Compound analysis via graph kernels incorporating chirality. *Journal of Bioinformatics and Computational Biology*, 8(1):63–81, 2010.
2. J. A. Castillo-Garit, Y. Marrero-Ponce, F. Torrens, and R. Rotondo. Atom-based stochastic and non-stochastic 3d-chiral bilinear indices and their applications to central chirality codification. *Journal of Molecular Graphics and Modelling*, 26(1):32–47, 2007.
3. B. Gaüzère, L. Brun, and D. Villemin. Two New Graphs Kernels in Chemoinformatics. *Pattern Recognition Letters*, 33(15):2038–2047, 2012.
4. P.-A. Grenier, L. Brun, and D. Villemin. Incorporating stereo information within the graph kernel framework. Technical report, CNRS UMR 6072 GREYC, 2013. http://hal.archives-ouvertes.fr/hal-00809066.
5. P.-A. Grenier, L. Brun, and D. Villemin. Incorporating molecules stereisomerism within the machine learning framework. In *Structural, Syntactic, and Statistical Pattern Recognition*, pages 12–21. Springer, 2014.
6. P.-A. Grenier, L. Brun, and D. Villemin. Taking into account interaction between stereocenters in a graph kernel framework. Technical report, CNRS UMR 6072 GREYC, 2014. https://hal.archives-ouvertes.fr/hal-01103318.
7. P.-A. Grenier, L. Brun, D. Villemin, et al. A graph kernel incorporating molecule's stereisomerism information. *Proceedings of ICPR 2014*, 2014.
8. J. Jacques, A. Collet, and S. Wilen. *Enantiomers, racemates, and resolutions.* Krieger Pub. Co., 1991.
9. H. Kashima, K. Tsuda, and A. Inokuchi. Marginalized kernels between labeled graphs. In *ICML*, volume 3, pages 321–328, 2003.
10. P. Mahé and J.-P. Vert. Graph kernels based on tree patterns for molecules. *Machine Learning*, 75(1):3–35, Oct. 2008.
11. N. Shervashidze, P. Schweitzer, E. J. Van Leeuwen, K. Mehlhorn, and K. M. Borgwardt. Weisfeiler-lehman graph kernels. *The Journal of Machine Learning Research*, 12:2539–2561, 2011.
12. M. Varma and B. R. Babu. More generality in efficient multiple kernel learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 1065–1072. ACM, 2009.
13. W. T. Wipke and T. M. Dyott. Stereochemically unique naming algorithm. *Journal of the American Chemical Society*, 96(15):4834–4842, 1974.