SYNERGETICS AND THEORY OF CHAOS

УДК 167.7, 517

Gritsak von Groener V.V., Gritsak-Groener J., Petrou M.

SUPPER ALGORITHM OF PATTERN RECOGNITION

Laboratory of HRIT Corporation, USA, UK
Royal Holloway, University of London

In this paper, we present a novel and efficient algorithm for solving direct and inverse problems that exist in chemical preparations with antivirial effects. As to the best of our knowledge, the proposed algorithm being reported in this paper is the fastest algorithm published in the literature. The algorithm requires that the pattern information/data to be provided as input in the form of a graph. Most other published algorithms do not have this as a requirement. The algorithm is based on a topos classifier of topos proximity space construction. This algorithm is very effective for pattern recognition problems with given geometric and numerical characteristics. The algorithm's effectiveness is illustrated by solving the problem of computing/finding the chemical molecule structure which is used in anti-AIDS preparations.

Key words: algorithm, pattern recognition, graph, topos.

1. Introduction

In this paper we propose an algorithm and methodology that computes:

a) Proximity measure between two objects represented as a graph in pseudomeasure inducing from all possible topologies and all the finite graph spaces (direct problem);
b) In all possible graphs space, the algorithm finds all possible graph objects that differ from a given graph by a distance of 0, 1, 2, … (pseudomeasure in integer calibration).

This problem was first introduced in [1[. A solution to it was recognized after the publication of article [2]. The work published in [1] and [2] were then used in applications such as pattern recognition, chemical design and computational methods in pharmacology [3].

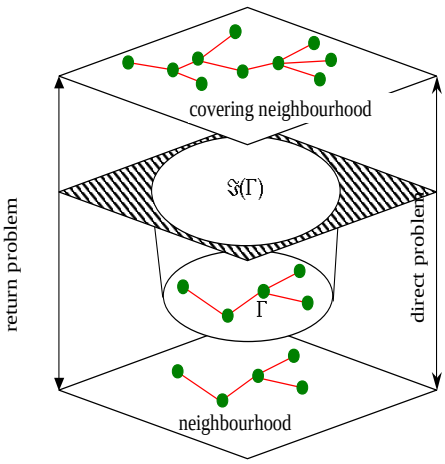

Fig. 1. The direct and return problem

It is important to note that for graph structures having a groupoid characteristics, the proposed algorithm and methodology have a wider use in a broader set of applications [4].

A number of biological and chemical systems, like two-sided coins, such as chemical structures, and viral infections are considered as statistical patterns. Such considerations do not correctly represent many important cases and applications.

It is well known, however, that the behavior of many atoms biochemical molecule and large population groups elements often is pliable to exact predications with help of single graph stereostructure taking account of graph kinetics.

Physical-chemical properties of the majority of chemical molecule and population depend on the structure of respective molecular graphs or graphs of mutually relations between population's element. This simple observation serves a basis for creation of fruitful and rapidly developing trend of theoretical chemistry and biology with a motto:

"structure – property".

A pharmacological method is the most direct way of struggle against illness virus. Unfortunately, this direction have been confronted with the great difficulties among them:

a) an extremely great computations volume connecting with the chemical structure choice for anti-virus preparations;

b) a big velocity of virus mutational variability (especially HIV-virus).

The above mentioned difficulties multiply the process of pharmacological preparations development. Therefore, a pharmacological chemist runs into a chosen solution/selection from several hundred millions of variants to optimal chemical structure and following "in vitro" test.

One of the interesting trends is the process of considering the description of chemical molecule similar to the given one by the totality of physical-chemical, pharmacological and other structure of the respective molecular graphs, conformations, hirality, etc.

Our suggested approach makes it possible to study random structures, the assigned structures and that assigned by the graphs, as to their structural proximity. The great number of such problems arise in the mathematical theory of pattern recognition and stochastic geometry.

2. Basic Idea

The main scheme of the Microscope Algorithm (MA) is based on the next propositions.

Suppose β is a *pattern* and α is an *image*. Consider the following steps of MA:

1) Using a distance concept of a topoi, if we get Distance(β, α) ~ 0, then the patters β is equals the image α.
2) Using a distance concept of a topoi, if we get Distance(β, α) > 0, then the patters β is not equals the image α.

Example 1

So, in a fig. 2 are shown the structure formulas: at the left – the amino acids triptofan ($C_{11}H_{12}N_2O_2$) and on the right – the medicine azidothymidine (AZT, $C_{10}H_{13}N_5O_4$), which have among themselves distance 0.

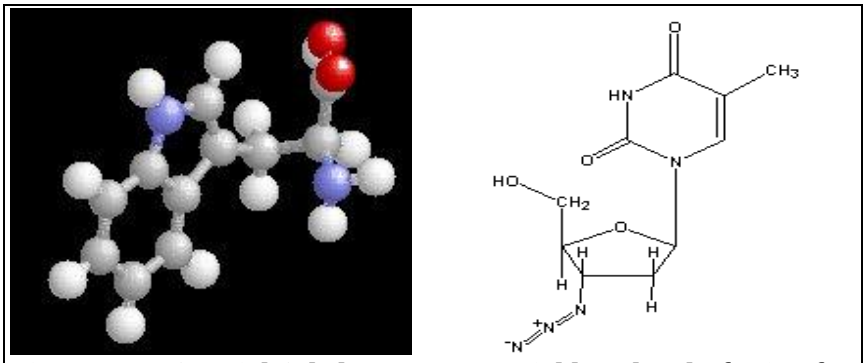

Fig. 2. AZT (on right) belongs to a 0-neighbourhood of triptofan (on left)

3. General Condition

Without loss of generality it can be assumed that:

1) $\forall$ objects is finite;
2) $\forall$ pattern is defined a *Input Information Graph* (IIG);
3) $\forall$ IIG is connected;
4) $\forall$ categories is defined of sketch (see [3]).

Definition. A topoi T is a category

$$T = (Mor(T), Ob(T)),$$

such that:

a) T is finitely complete and cocomplete;
b) T has exponential.
5) (For every a, b $\in$ Ob(T) exist object b[[a]] $\in$ Ob(T) and morphism

$$ev: b[[a]] \times a \longrightarrow b$$

is commutative).

c) T has a subobject classificatory.

(For every monomorphism f: a $\longrightarrow$ d, f $\in$ Mor(T), exists every one morphism

$$\chi[f]: d \longrightarrow \Omega$$

is Cartesian square).

The object Ω(T) is name a classificatory of topoi T.

4. Foundation Principles and Theorem 1

Topoi Principle

A classificatory Ω(T) is full defined every structural property of a topoi T.

Theorem 1

Let $\Im$ is a topological space and B($\Im$) is the category of all bundle for $\Im$; then B($\Im$) is the topoi.

5. Step 1

Definition

Let a graph Γ = (V, W) is IIG of a pattern 🐅, where V is the set of vertices Γ and W is the set of edges Γ. We constructed the new graph Ξ (Γ) = (V_1, W_1). Ξ(Γ) is called *deltoid* Ξ(Γ) of a graph Γ. V_1 is a disjunctive union of vertices V, i.e. (V_1 = V$\coprod$V). An edge w_d = (v_1, v^2)∈W_1 if and only if an edge w = (v_1, v_2)∈W, where v_1∈V and the vertex v^2∈V_1 is analogue the vertex v_2∈V in the second component V_1.

Step 1.

We constructed the deltoid Ξ(Γ) of the graph Γ = (V, W), where Γ is IIG of the pattern 🐅.

Example 2

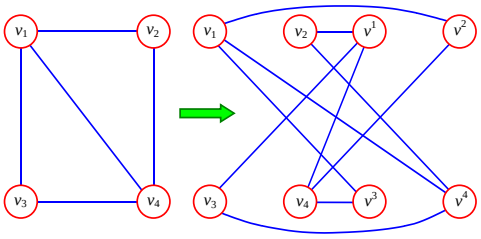

Fig. 3. The graph Γ (on left) and the deltoid Ξ(Γ) (on right)

6. Step 2

Theorem 2.

Suppose Ξ(Γ) is a deltoid of graph Γ. Then there are two transitive ordering (≥ and ⋉) of Ξ(Γ), only.

Step 2.

We constructed the transitive ordering deltoids $\Xi^{\geq}$(Γ) and $\Xi^{\ltimes}$(Γ).

7. Step 3

Theorem 3

Suppose $\Xi^{\geq}$(Γ) (or $\Xi^{\ltimes}$(Γ)) is the graph with in Step 2. Then there is only one finite topological space $\Im(\Xi^{\geq}(\Gamma))$ (or $\Im(\Xi^{\ltimes}(\Gamma))$).

Scheme of subalgorithm for step3

Let $\Xi^{\geq}$(Γ) = (V, E, ≥) is a transitive-order graph in step 2. $\Im(\Xi^{\geq}(\Gamma))$ = (V, $\Im$(V)) is the topological space, where V is the groud set of $\Xi^{\geq}$(Γ) and $\Im$(V) is a open subsets of $\Xi^{\geq}$(Γ). Here $\Im$(V) is induced by the subbasic subsets B(V) = {v ∪ {v_i}}, where v ∈V, v_i ∈V are all elements V incidental to the element v.

The topological space $\Im(\Xi^{\ltimes}(\Gamma))$ is constructed analogously.

Step 3.

We constructed the topological spaces $\Im(\Xi^{\geq}(\Gamma))$ and $\Im(\Xi^{\asymp}(\Gamma))$, respectively.

8. Step 4

Corollary of Theorem 1

If the $\Im(\Xi^{\geq}(\Gamma))$, $\Im(\Xi^{\asymp}(\Gamma))$ topological space defined in step 3, then there are topoi

$$\Im o\rho_1 = \Im o\rho\ (\ \Im(\ \Xi^{\geq}(\Gamma)),\ \Theta^{\geq}\ ),\ \Im o\rho_2 = \Im o\rho\ (\ \Im(\ \Xi^{\asymp}(\Gamma)),\ \Theta^{\asymp}\ ),$$

where $\Theta^{\geq}$ and $\Theta^{\asymp}$ are the classificatory of topoi $\Im o\rho_1$ and $\Im o\rho_2$, respectively.

Step 4

We constructed the topological spaces $\Im o\rho_1(\Gamma) = \Im o\rho\ (\Im(\Xi^{\geq}(\Gamma)),\ \Theta^{\geq})$ and $\Im o\rho_2(\Gamma) = \Im o\rho(\Im(\Xi^{\asymp}(\Gamma)),\ \Theta^{\asymp}\ )$, respectively.

9. Step 5

The following theorems are needed for the step 5.

Theorem 4

Suppose $\Gamma\square(\geq)$ is a finite transitive-ordering digraph, then there is the pseudometrical topological spaces $\Im(\Gamma\ \square(\geq))$, with a strongest pseudometric d: $\Im(\Gamma\ \square(\geq))\times\Im(\Gamma\ \square(\geq)) \longrightarrow Z^{+}$ and d is unambiguous.

Theorem 5

Suppose T = (Mor(T), Ob(T)) is topoi, then there is the digraph $I\eta(T) = (V_T, E_T)$, where V_T = Ob(T) is the set of the vertex $I\eta(T)$ and E_T is the set of arrows $I\eta(T)$. The digraph $I\eta(T)$ is called *inflation graph* of the topoi T. The inflation digraph $I\eta(T)$ is transitive ordering.

Corollary of theorem 4 and 5

Suppose $\Im o\rho_1(\Gamma) = \Im o\rho(\Im(\Xi^{\geq}(\Gamma)),\ \Theta^{\geq})$ and $\Im o\rho_2(\Gamma) = \Im o\rho(\Im(\ \Xi^{\asymp}(\Gamma)),\ \Theta^{\asymp})$ are topoi determined by step 4, then there are the pseudometrical topological spaces

$$\Im_1(\Gamma) = \Im_1(I\eta(\Im o\rho_1)) \text{ and } \Im_2(\Gamma) = \Im_2(I\eta(\Im o\rho_2)).$$

Step 5

We constructed the pseudometrical topological spaces $\Im_1(\Gamma)$ and $\Im_2(\Gamma)$.

10. Steps 6-10

Step 6 - Step 10.

Inverse transformation the detailed see in [6] and [7].

We constructed the topological spaces $\Im_1(\text{؞})$ and $\Im_2(\text{؞})$ of topoi $\Im o\rho_1(\text{؞})$ and $\Im o\rho_2(\text{؞})$, respectively. ؞ is the image of pattern recognition in the IIG form.

11. Step 11 – The Last Step

Theorem 6

Since the step4, we obtain the finite pseudometrical topological spaces $\Im^1(\Im^{\geq}(\Gamma))$ and $\Im^2(\Im^{\blacktriangleright}(\Gamma))$, where $\Im^{\geq}(\Gamma)$ and $\Im^{\blacktriangleright}(\Gamma)$ is the transitive ordering graphs. $\Im^1(\Im^{\geq}(\Gamma))$ and $\Im^2(\Im^{\blacktriangleright}(\Gamma))$ has strongest psevdometric, then exist the topological pseudometrical space:

$$\Im_{\geq\geq}(\Gamma) = \Im^1(\Im^{\geq}(\Gamma)) \times \Im^2(\Im^{\blacktriangleright}(\Gamma)). \qquad (1)$$

$\Im_{\geq,\geq}(\Gamma)$ is the Cartesian product topological spaces of $\Im^1(\Im^{\geq}(\Gamma))$ and $\Im^2(\Im'(\Gamma))$ and $\Im_{\geq,\geq}(\Gamma)$ have the strongest pseudometric. $\Im_{\geq,\geq}(\Gamma)$ is called a *Structure Distanced Space* of *Pattern Recognition.*

Step 11

Constructed the topological pseudometrical space:

$$\Im_{11} = \Im_1((\Gamma)) \times \Im_1(\text{♂}) \text{ and } \Im_{22} = \Im_2(\Gamma) \times \Im_2(\text{♂}).$$

Final

Finally, we get

$$d1 = \text{distance } ((\Im^{\geq}(\text{♀}) \times \text{int}^{\geq}(\text{♀})), \Im'(\text{♂}) \times \text{int}^{\geq}(\text{♂})),$$

$$d2 = \text{distance } ((\Im^{\times}(\text{♀}) \times \text{int}^{\times}(\text{♀})), \Im^{\times}(\text{♂}) \times \text{int}^{\times}(\text{♂})),$$

where $\Im^{\geq}(\text{♀})$, $\Im^{\times}(\text{♀})$, $\Im'(\text{♂})$, $\Im^{\times}(\text{♂})$ - the classificatory objects of a pattern ♀ and an image ♂, and $\text{int}^{\geq}(\text{♀})$, $\text{int}^{\times}(\text{♀})$, $\text{int}'(\text{♂})$, $\text{int}^{\times}(\text{♂})$ - initial objects of a pattern ♀ and an image ♂. Then

$$d = d1 + d2 \qquad (2)$$

is distance of a pattern ♀ and an image ♂. d from (2) is called an *Absolute Distanced* between a pattern ♀ and an image ♂.

12. Example 3

Fig. 3 shows the 4 classical nucleotides: adenine, cytosine, guanine, uracile, and in the middle - AZT, for which graph of the formula is on a minimum total distance from them.

12. Conclusion

In this paper, we presented a novel and efficient algorithm for solving direct and inverse problems that exist in chemical preparations with antivirial effects. As to the best of our knowledge, the proposed algorithm reported in this paper is the fastest algorithm published in the literature. We illustrated the algorithm with some examples demonstrating that the algorithm is very effective for pattern recognition problems with given geometric and numerical characteristics. This project is still at its infancy. We plan to report more details about this algorithm and applications that could be impacted as a result.

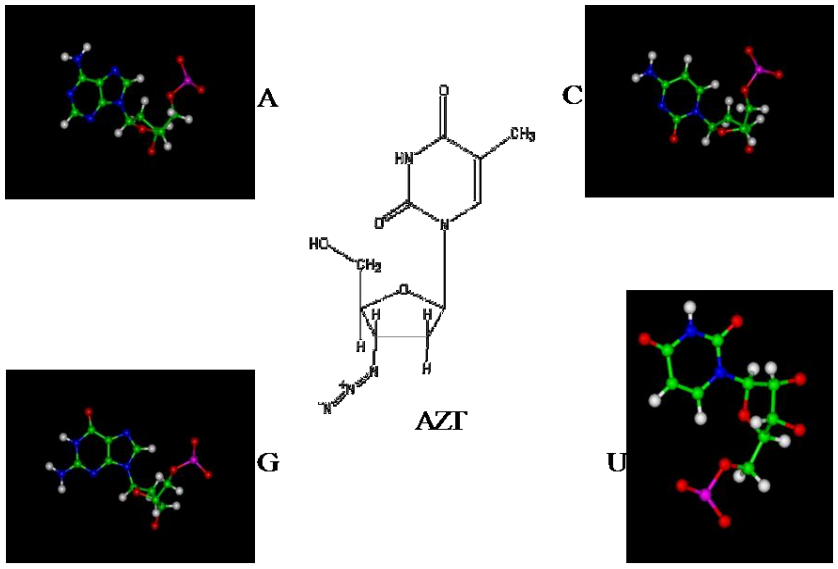

Fig. 4. AZT (on middle) has a minimum total distance from the nucleotides A, C, G and U

R e f e r e n c e s :

1. *Gritsak–Groener V. V.* The Computing Proximity of Chemical Compound with the Help of Topoi-Categorical Analysis, Proceeding of two USSR Interinstitute Conference "The Molecular Graphs in Chemical Researches". Kalinin, (1990)5.
2. *Gritsak–Groener V. V.* Proc. Nat. Acad. Sci. Ukraine, ser. A, N3(1993)100.
3. *Gritsak–Groener V. V.* Fundamental of Mathematical Cybernetics, 2004.
4. *Gritsak–Groener V. V., Gritsak–Groener J*. Gruppoid-Algorithm of Pattern Recognition // Proceedings of the XX-th IUPAP International Conference on Statistical Physics (STATPHYS"20). — Paris, 1998.
5. *Gritsak–Groener V. V., Gritsak–Groener J*. Pattern Recognition Algorithms in Statistical Physics. // The Proceedings of the XX-th IUPAP International Conference on Statistical Physics (STATPHYS"20). — Paris, 1998.
6. *Gritsak–Groener V. V., Gritsak–Groener J*. ARTS COMBINATORIA. — Charkiv, NTU “ChPI”, 2003.

Статья поступила в редакцию 10.02.2010 г.

Гритсак-Грёнер В.В., Гритсак-Грёнер Ю., Петроу М.

Супералгоритм распознавания образов

Представлен новый и эффективный алгоритм для решения прямых и обратных задач, которые возникают в химических процессах с антивирусными эффектами. Насколько нам известно из литературы, предлагаемый алгоритм является самым быстрым алгоритмом. Алгоритм требует, чтобы исходная информация была представлена в форме диаграммы. Большинство других опубликованных алгоритмов не предъявляют этого требования. Алгоритм основан на построении топос-классификатора. Этот алгоритм очень эффективен для задач распознавания образов с геометрическими и числовыми данными. Эффективность алгоритма проиллюстрирована решением задачи построения химической структуры молекулы, которая используется в реакциях анти-СПИД.

Ключевые слова: алгоритм, распознавание изображений, диаграмма, топос.