Limit combinatorics as a method for investigating of mathematical models

Vladimir Remeslennikov
Sobolev Institute of Mathematics, Siberian Branch of Russian Academy of Science, Omsk, Russia
E-mail: remesl@ofim.oscsbras.ru

Abstract. In recent years, due to the applications in theoretical informatics, the researches on finite combinatorics (graphs, directed graphs, partial orders, and so on) and the applications of these studies in the economy, in the social sphere and in statistical physics have been intensively developing. The main objective is to present the latest mathematical studies on formal languages, large networks, finite combinatorics and model theory to a wide range of researchers involved in the development of economic and social models, information security models, the growth of the Internet models, and so on, for whom, by the author’s opinion, such the studies would be useful.

1. Introduction
The article is on the popular science side. The main objective is to present the latest mathematical studies on formal languages, large networks, finite combinatorics and model theory to a wide range of researchers involved in the development of economic and social models, information security models, the growth of the Internet models, and so on, for whom, by the author’s opinion, such the studies would be useful. For specific practical tasks there are the complicated working models to be built up, while our current task is to show them in the very core of the subject, to reveal their unification mathematical idea. It would be naive to suppose that our recommendations can satisfy the needs of any and all mathematical models developers. This is not so, and we formulate the requirements that are imposed on the mathematical models. They are the following:

(i) the model must be dynamical in discrete time, i.e. must consist of a finite large (or infinite) number of finite combinatorics

\[ C_1 \rightarrow C_2 \rightarrow \ldots \rightarrow C_k \rightarrow \ldots \; \]  \hspace{1cm} (1)

(ii) it should be described in formal mathematical language transitions

\[ \alpha_i : C_i \rightarrow C_{i+1} . \]

Good examples of such models are currently served by different models of Internet growth [1].

For the models above, the notion of a \( \Delta \)-index (this is a rational number) is defined, as well as the suitable \( \Delta \)-properties, where \( \Delta \) is a concrete finite combinatorics. The main mathematical problems here are the following two mathematical problems having good practical applications:
(i) Is the combinatorics sequence (1) convergent, and how does its limit appear $C_\infty$?

(ii) Which $\Delta$-properties are quasi-random, which means their stability with respect to the passage to $C_\infty$? (The exact definition, see, for example, in [12,13]).

The universality of the mathematical approach to a wide range of such models is achieved, first, by abstracting to arbitrary categories, and secondly, by treating objects of categories as models of a certain language, which allows describing equations, inequalities, and so on in them. Let us explain in more detail what is being discussed.

2. The category of combinatorics

The definition of the category theory can be found in many sources. For readers convenience, we provide the definition from Wikipedia (in Russian): “The category theory is a section of mathematics that studies the properties of relations between mathematical objects that do not depend on the internal structure of objects”. More formally, a category $\mathcal{C}$ consists of the following three mathematical entities:

1. a class of objects $\mathcal{C}$;
2. for every pair of objects $A$ and $B$ it is defined a set of morphisms $\text{Hom}(A, B)$;
3. for every pair of morphisms $f \in \text{Hom}(A, B)$ and $g \in \text{Hom}(B, C)$ it is defined a composition $f \circ g \in \text{Hom}(A, C)$;
4. for each object $A$ it is defined an identity morphism $\text{id}_A \in \text{Hom}(A, A)$,

such that the following two axioms hold:

1. associativity: $h \circ (g \circ f) = (h \circ g) \circ f$;
2. identity: $f \circ \text{id}_A = \text{id}_B \circ f = f$ for all $f \in \text{Hom}(A, B)$.

Let us list the most important examples of categories in the study of mathematical models.

- The category of sets. Objects here are sets and morphisms are maps between sets.
- The category of simple graphs. Objects here are simple graphs (i.e., undirected graphs in which both multiple edges and loops are disallowed) and morphisms are maps which preserve edges.
- The category of partial orders. The partial order is a set with a binary relation $\leq$ with the following axioms:
  1. reflexivity: $\forall a \ a \leq a$;
  2. transitivity: $\forall a, b, c \ (a \leq b) \land (b \leq c) \rightarrow (a \leq c)$;
  3. antisymmetry: $\forall a, b \ (a \leq b) \land (b \leq a) \rightarrow (a = b)$;

  Morphisms in this category are maps preserving the order relation.

3. The language of equations and inequalities for combinatorics

The concepts of the equation and inequalities are well known for fields, including the field of real and complex numbers. In the monograph of three authors E.Yu. Daniyarova, A.G. Myasnikov and V.N. Remeslennikov “Algebraic geometry over algebraic structures” [2], the concepts of equation and inequality for combinatorics of all types are introduced, if a language is defined for them. Here the language is understood as those relations (predicates) and functions that are defined for the class of combinatorics $\mathcal{C}$. For example, for the category of simple graphs, the language will consist of two predicates — the vertex equality predicate and the neighborhood predicate for two vertices $E(x, y)$. For these predicates, the following natural axioms are fulfilled: $\neg E(x, x)$ (hence, there are no loops in the graph), $E(x, y) \iff E(y, x)$ (in other words, this predicate is symmetric with respect to the pair of vertices). For the category of partial orders, the language
will consist of predicates of equality and a predicate of order $\leq$, and the axioms for it are given in the previous section. The monograph [2] explains what is an equation and an inequality in this case, and how to solve such systems of equations and inequalities.

We also note that many properties of simple graphs can be written in this language. For example, the fact that in a finite graph there are 10 different triangles or there are no pentagons and so on.

Let us explain this more formally. Let $\Delta$ be a fixed finite simple graph with $n$ vertices and $\Gamma$ be another finite simple graph. It is not hard to write a formula in the language $L$ that states “graph $\Gamma$ has hereditary subgraph $\Delta$”. Let $n = 3$ and $\Delta$ be the triangle, then the desired formula is of the form: $\exists x_1, x_2, x_3 (x_1 \neq x_2 \land x_1 \neq x_3 \land x_2 \neq x_3 \land E(x_1, x_2) \land E(x_1, x_3) \land E(x_2, x_3))$.

Hence it is clear that we can write down the formula by the fact that $\Gamma$ contains exactly $k$ of different triangles. If $k$ is the maximum number contained in $\Gamma$ triangles, then we say that the $\Delta$-index of the graph $\Gamma$ is $k$.

4. Quasi-random combinatorics

In many cases, the finite combinatorics must be digitized either by elements from the field or by elements of some Boolean algebra. The notion of a random graph and a quasi-random graph arise on this path. Let us give these definitions following the classical books on the graph theory.

We start with the definition of a quasi-random graph. By definition, a quasi-random graph on $n$ vertices is a graph $\Gamma$ satisfying the following properties:

(i) $\Gamma$ is a complete graph;
(ii) to each edge on the vertices $x, y$ the probability of appearance of this edge in the graph $\Gamma$ is assigned. For example, if this probability is 0, then this means that corresponding edge actually absent in the graph $\Gamma$.

A quasi-random graph $\Gamma$ is called a random graph if the real number $p \in (0, 1]$ is chosen and the probability of occurrence of each edge in the graph $\Gamma$ is equal to the number $p$.

5. Large networks and $\Delta$-indexes

In this report, we will only work with so-called dynamic large networks. Let us explain what this means. We have a finite or infinite sequence of finite combinatorics:

$$C_1 \to C_2 \to \ldots \to C_k \to \ldots,$$

where $C_k$ are graphs (their nature can be different: quasi-random graphs, multi-graphs and so on), symbol “$\to$” means a morphism in this category, and the number of elements in combinatorics $C_k$ increases asymptotically to infinity.

Let us list the main tasks associated with such sequences.

(i) It is necessary to correctly define the notion of convergence with the help of $\Delta$-indexes.
(ii) Determine when the limit combinatorics $C_{\infty}$ exists.
(iii) If it possible, determine $\Delta$-index for the limit combinatorics $C_{\infty}$.
(iv) Determine what properties of this sequence are quasi-random.

Over the past 10 years, many mathematicians dealt with these problems. Also there have been published a large number of works in this area, many of which are with a practical trend [3–14].
6. Limit combinatorics

Let us define one useful metric on graphs, which occurs in many mathematical models. Let $H$ and $G$ be two simple graphs. Assume that the set of vertices $V(H)$ of the graph $H$ have cardinality $m$, and $|V(G)| = n$, $m \leq n$. Suppose that an embedding $\alpha : V(H) \rightarrow V(G)$ is given.

We define the metric $ED(\alpha)$ as the number of errors in the embedding $\alpha$. We consider an error to be the mapping of incident vertices into nonincident vertices or the mapping of nonincident vertices into incidence vertices. The edit distance between graphs $H$ and $G$ is the following value:

$$ED(H, G) = \min_{\alpha} ED(\alpha)/C_n^2.$$  

With this distance, the sequence (1) turns into a metric space, and some properties of combinatorics can be extracted with the help of topological arguments.

7. Statistical approach to research

Let us define the central concept of the statistical method for research of combinatorics. This is the concept of the density of one combinatorics relative to the other one. It allows us to enter finite combinatorics quasi-random. Particularly useful are the combinatorics exchange indices. This is a fairly rare case. Often for practical applications, it is important that convergence is relative to the most important properties that are defined by the combinatorics $\Delta_1, \Delta_2, \ldots, \Delta_k$. If we talk about stock exchanges, then there is a finite set of main exchange indices.

Let $\Delta$ and $G$ be two simple graphs, $|V(\Delta)| = m$, $|V(G)| = n$, $m \leq n$. Denote by $\rho(\Delta, G)$ the number of induced subgraphs in $G$, that are isomorphic to $\Delta$. Then denote the density of the combinatorics $\Delta$ in $G$ by $\rho^*(\Delta, G) = \frac{\rho(\Delta, G)}{C_m^n}$, where $C_m^n$ is the binomial coefficient. Further, fixing a graph $\Delta$ and using the sequence (1) one can construct the following sequence of rational numbers:

$$\rho^*(\Delta, C_1), \rho^*(\Delta, C_2), \ldots, \rho^*(\Delta, C_k), \ldots$$  

**Definition.** The sequence (1) is called convergent, if for any $\Delta$ the sequence (2) is a convergent rational sequence.

This definition is very rigid, and it is often not fulfilled. Fortunately, the research customer is only interested in the density with respect to a finite number of properties $\Delta = \{\Delta_1, \ldots, \Delta_k\}$, so we give a $\Delta$-variant of this definition.

**Definition.** The sequence (1) is called $\Delta$-convergent, if the sequence (2) is a convergent rational sequence for all selected combinatorics $\Delta_1, \ldots, \Delta_k$.

References

[1] Raygorodskiy A M 2010 Models of random graphs and its applications Proceedings of MIPT 4(2) pp 130–140
[2] Daniyarova E Yu, Myasnikov A G and Remeslennikov V N 2016 Algebraic geometry over algebraic structures (Novosibirsk: SO RAN) 243 p.
[3] Chung F and Graham R 1991 Quasi-random tournaments Journal of Graph Theory 15(2) pp 173–198
[4] Chung F, Graham R and Wilson R 1989 Quasi-random graphs Combinatorica 9(4) pp 345–362
[5] Diaconis P and Janson S 2008 Graph limits and exchangeable random graphs Rendiconti di Matematica, Serie, VII 28 pp 33–61
[6] Erdős P 1962 On the number of complete subgraphs contained in certain graphs Publ. Math. Inst. Hungar. Acad. Sci. 7 pp 459–464
[7] Griffiths S 2013 Quasi-random oriented graphs Journal of Graph Theory 74(2) pp 198–209
[8] Jagger C, St’ovic P and Thomason A 1996 Multiplicities of subgraphs Combinatorica 16(1) pp 123–141
[9] Krivelevich M and Sudakov B 2006 Pseudo-random graphs In: More Sets, Graphs and Numbers. Bolyai Society Mathematical Studies 15 (Spring-Verlag) pp 199–262
[10] Razborov A 2007 Flag algebras Journal of Symbolic Logic 72(4) pp 1239–1282
[11] Razborov A 2013 On the Caccetta-Haggkvist conjecture with forbidden subgraphs Journal of Graph Theory 74(2) pp 236–248
[12] Razborov A 2013 What is a Flag Algebra? Notices of the AMS 60(10) pp 1324–1327
[13] Coregliano L and Razborov A On the Density of Transitive Tournaments arXiv:1501.04074
[14] Lovasz L 2012 Large networks and graph limits AMS Colloquium publications 60 pp 473