Constructive enumeration in AGT: goals, techniques, history and examples

Mikhail Klin
(Ben-Gurion University and Matej Bel University)

September 15, 2015
1. Enumeration in combinatorics: general concepts
2. Constructive enumeration
3. Constructive enumeration of graphs
4. Enumeration of incidence systems
5. Other approaches to constructive enumeration
6. Constructive enumeration of association schemes
7. From COCO to COCO II
8. Some fulfilled projects
9. References
We provide a very naïve introduction to this area of research.

No attempt to reach comprehensive covering of any of the ingredients, mentioned in the title.

Our main interest is to prepare the audience to follow in the further lectures to a few concrete champion achievements in constructive enumeration.

Special attention to the approach originated by Igor Faradžev.
Graphs and incidence structures are two main kinds of structures.

Description by lists, matrices (adjacency or incidence), diagrams.

Mostly will be concentrated on graphs, in most cases simple.

A graph has \( n \) vertices, \( m \) edges, the valency \( v(x) \) of a vertex \( x \in V \) is the number of its neighbours.
The graphs $\Gamma = (V, E)$ and $\Gamma' = (V', E')$ are called \textbf{isomorphic} if there exists a bijection $\pi : V \rightarrow V'$ which induces a bijection between $E$ and $E'$.

In other words, for each $\{x, y\} \in E$ the edge $\{x^\pi, y^\pi\} \in E'$.

The set $\text{Aut}(\Gamma)$ of all isomorphisms of $\Gamma$ with itself forms a group.

Different graphs, isomorphic to $\Gamma$, correspond to cosets of $\text{Aut}(\Gamma)$ in $\text{Sym}(V)$. 
Enumeration of graphs (with $n$ vertices) means knowledge of all isomorphism classes of $n$-vertex graphs.

We distinguish between analytical and constructive enumeration.

Analytical enumeration results in a generating function. Sometimes it is called counting.

Constructive enumeration should provide the set of representatives (transversal) of all isomorphism classes.
The problem can be reformulated in group-theoretical terms.

Start from the symmetric group $S_n$ of degree $n$, acting on the set $[1, n] = \{1, 2, \ldots, n\}$.

Consider its induced action $S_n^{[2]}$ on the set $\binom{[1,n]}{2}$ of 2-subsets of $[1, n]$.

For each $0 \leq m \leq \binom{n}{2}$ determine the orbits of $S_n^{[2]}$ on the set of $m$-subsets of $\binom{[1,n]}{2}$. 
Example (2.1 Graphs with 4 vertices)

a) Analytical enumeration

\[ Z(S_4, [1, 4]) = \frac{1}{24} (x_1^4 + 6x_1^2x_2 + 3x_2^2 + 8x_1x_3 + 6x_4) \]

\[ Z(S_4^{[2]}) = \frac{1}{24} (x_1^6 + 9x_1^2x_2^2 + 8x_3^2 + 6x_2x_4) \]

\[ f(x) = f(S_4^{[2]}, 1 + x) \]

\[ = \frac{1}{24} ((1 + x)^6 + 9(1 + x)^2(1 + x^2)^2 \]

\[ + 8(1 + x^3)^2 + 6(1 + x^2)(1 + x^4)) \]

\[ = 1 + x + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6. \]

In particular,

\[ f(x) \big|_{x=1} = \frac{1}{24} \left( 2^6 + 9 \cdot 2^4 + 8 \cdot 2^2 + 6 \cdot 2^2 \right) \]

\[ = \frac{2^3}{24} (8 + 18 + 4 + 3) = \frac{33}{3} = 11. \]
b) Constructive enumeration

I am using diagrams of graphs without labels of vertices. (Compare with the smile of Cheshire cat which disappears.)

We speak of abstract graphs.
Practically we need to introduce canonical labelling of graphs from the same isomorphism class.

Classical approach via matrices aka lexicographical order on 2-subsets.

Consider maximal (or minimal) upper triangle of the adjacency matrix.
Example (2.2)

- Abstract graph $\Gamma$ with 4 vertices and 5 edges.
- $|Aut(\Gamma)| = 4$.
- Thus there are $\frac{4!}{4} = 6$ concrete labelled graphs in the isomorphism class.
This is the canonical labelling: $A(\Gamma) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 \end{pmatrix}$. 
Example (2.3)

- Block design $D$ with $\nu = 6$ points, $b = 10$ blocks of size $k = 3$; each point appears in $r = 5$ blocks, each pair appears in $\lambda = 2$ blocks.
- We also wish to get its canonical incidence matrix.
- We list it as set of blocks.
Example (2.3 cont.)

- $D = (P, B)$.
- $B = \{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 5\}, \{1, 4, 6\}, \{1, 5, 6\}, \{2, 3, 6\}, \{2, 4, 5\}, \{2, 5, 6\}, \{3, 4, 5\}, \{3, 4, 6\}\}$.
- We will return to this example later on.
Formulation of the problem: Describe an algorithm providing a transversal for the set of all graphs with prescribed properties.

Such a process is usually called an exhaustive search.

Opposite to heuristical approach or to random algorithms.
We distinguish three main strategies (due to McKay):

- Read-Faradžev type algorithms;
- McKay type algorithms;
- Group-theoretical algorithms, based on the use of double cosets and the so-called homomorphism principle.

We will mainly concentrate on the first strategy.

The other types will be mentioned briefly.
Example (2.4)

Connected regular graphs on 8 vertices of valency 3 (cubic graphs)
All are Hamiltonian.
Orders of $Aut(\Gamma)$ are 48, 16, 2, 4, 4.
First approach: generate canonical matrices step by step.

Second approach: First look on graph theoretical properties
- bipartite, girth 4;
- non-bipartite, girth 4;
- 3, 4, 2 triangles, girth 3.

Third approach: investigate the interplay between $D_8$ of order 16 and $S_4 \wr S_2$ of order $4!2^4 = 384$ (provided that we are aware that all graphs are Hamiltonian).
Example (2.5 Construction of all isomers of dioxine)

- Excursion to mathematical chemistry, especially to the Bayreuth school (Kerber, Laue et al).
- Start with skeleton $\Delta$ of dioxine: $C_{12}O_2H_4Cl_4$
There are 8 remaining positions.
4 of them may be occupied by $H$, 4 by $Cl$.
All possibilities are allowed.
Describe all non-equivalent isomers.
First, let us use the Cauchy-Frobenius-Burnside Lemma (CFB).

$\text{Aut}(\Delta) \cong E_4$ (Klein group).

We consider the action of $E_4$ on the 8 empty positions.

The number of orbits on $\binom{[1,8]}{4}$:

$$t = \frac{1}{4} \left( \binom{8}{4} + 3 \cdot \binom{4}{2} \right) = \frac{70 + 3 \cdot 6}{4} = \frac{88}{4} = 22.$$ 

There are 22 equivalency classes.

The result of the constructive enumeration can be found at page 73 of KerL98.
Here, constructive enumeration is fulfilled in terms of group theory.

Consider the pair \((A, B)\) of groups, \(A = E_4\), \(B = S_4 \times S_4\), both as subgroups in \(S_8\).

The problem is reduced to the enumeration of double cosets.

Additional reduction to double cosets in \(S_4\).

For details see KerL98.
Canonical labelling is again crucial.

Two graphs $\Gamma$ and $\Gamma'$ are isomorphic
\[\iff \text{Can}(\Gamma) = \text{Can}(\Gamma').\]

We need to construct new and new graphs and to reach their canonical isomorphs.

Read-Faradžev approach is based on lexicographical ordering.

It disregards naive benefits from manipulation with local invariants.
Example (2.6 Graphs with 5 vertices and 5 edges via Read-Faradžev approach)

- Solutions (with valencies):

  - (2^5)
  - (3, 2^3, 1)
  - (4, 2^2, 1^2)
  - (3, 2^3, 1)

- Each set of valencies is considered separately.
- Let us deal with (3, 2^3, 1).
- We disregard the difference between the lengths of cycles.
• Each time we remember the information about the stabilizer of the considered partition.
• We continue with those solutions which have a chance to be extended to full solutions.
• Partial valencies are respected.
• We got three solutions; the first and third provide canonical isomorphs.
A few helpful tricks for pruning the search tree.

- Removing non-graphical valency sequences for partial solutions;
- Manipulation with permutations of graphical sequences;
- Use of semicanonical labelling (construction of graphs is done row by row);
- Backtrack canonicity check;
usage of found automorphism of generated subgraphs to cut branches;
prediction of zeros in further levels;
using the proof of non-canonicity to skip other non-canonical solutions;
cancelling unnecessary canonicity check on the basis of canonicity of the previous graph;
forcing.
McKay’s approach

- Of course he also uses canonical isomorphs and many ideas, like in previous approach.
- However, his way to get a canonicity check is essentially different: no relation to lexicographical order.
- In a sense, the result of producing the canonical isomorph appears as the output of a certain algorithm.
- It is based on the use of equitable partitions (tactical decompositions).
Example (2.7)

Start with a graph

\[
\begin{array}{ccc}
  & 1 & 2 \\
3 &   & 4 \\
  & 5 & 6 \\
\end{array}
\]
Example (2.7)

Distinguish vertices of valency 4, get equitable partition.
Example (2.7)

Distinguish one more vertex, get again equitable partition.
Example (2.7)

Distinguish one more vertex.
Example (2.7)

- All vertices are distinguished.
- Now order them in a certain way (for details, read carefully Brendan’s texts).
- Get canonical isomorph.
- This is a very simplified explanation of McKay’s approach to canonisation.
• Generation procedures in FR and McK approaches are more similar (than canonisation approaches).
• However, here McKay pays attention to the use of local invariants of graphs.
• In general he is interested in graphs with some hereditary properties.
- Both approaches are working with the automorphism group of the graphs.
- In both cases the group $\text{Aut}(\Gamma)$ of the graph under consideration appears as part of the result.
- Nevertheless there is some evidence to claim that for small groups, McKay’s approach is more efficient.
- nauty $= (\text{No AUTomorphisms, Yes?})$
- Do no substitute ‘!’ for ‘?’.
We touch this issue very briefly.

Canonical incidence matrix should be defined: consider the direct product of symmetric groups, acting on points and blocks.

Faradžev and A.V.Ivanov elaborated a very efficient orderly algorithm.

Nowadays the champions in isomorph-free generation of designs are P. Kaski and P. Östergård.
Example (2.3 revisited)

- \((v, b, k, r, \lambda) = (6, 10, 3, 5, 2)\)
- Search tree below:
• **Kramer-Mesner** method is another famous approach.

• Assume we are looking for a design $D = (P, B)$ with the parameters $(v, b, k, r, \lambda)$.

• Assume there is evidence to think that $\exists D$ which allows a suitable permutation group $(H, P)$ as a subgroup of $Aut(D)$.

• Then the method outlined below may work quite efficiently.
- Describe all orbits of \((H, P)\) on \(k\)-subsets and 2-subsets of \(P\).
- Create "inclusion" matrix.
- Manipulate by columns, watching rows.
- All solutions are suitable subsets of columns.
- Of course, check which solutions are isomorphic.
Example (2.3 once more)

- We are looking for \((6, 10, 3, 5, 2)\) design.
- Consider the "point residual" structure for any point, say 1.
- Observe that in our case this is a regular graph of valency 2 on 5 points.
- Up to isomorphism only the pentagon \(P = C_5\) exists.
\begin{itemize}
    \item \( \text{Aut}(P) = D_5 \), a group of order 10.
    \item Assume that the design is invariant with respect to \( D_5 \). Here, \( D_5 \) acts intransitively on \([1, 6]\).
    \item Start with an arbitrary copy of the cycle \( C_5 \) on \([2, 6]\).
    \item \( D_5 = \langle (2, 3, 4, 5, 6), (3, 6)(4, 5) \rangle \).
    \item Representatives of orbits:
        \begin{itemize}
            \item 12 – 23 – 24
            \item 123 – 124 – 234 – 235.
        \end{itemize}
    \item KM-matrix:
\end{itemize}

\[
\begin{array}{cccc}
    & A & B & C & D \\
\hline
123 & 12 & 124 & 234 & 235 \\
12 & 12 & 2 & 2 & 0 & 0 \\
23 & 1 & 0 & 2 & 1 \\
24 & 0 & 1 & 1 & 2 \\
\end{array}
\]
We get two solutions: $A \cup D$ and $B \cup C$. They are isomorphic to the canonical one. In fact, $\text{Aut}(D_5) \cong A_5$. 
This approach works also for $t$-designs with $t > 2$.
If the group $H$ is large, the description of $k$-orbits and $t$-orbits may be quite sophisticated.
Efficient technique is elaborated at Bayreuth.
A lot of $t$-designs were constructed.
Building kit paradigm may be formulated as follows:

- the desired combinatorial structure \( \gamma \) may be obtained via suitable manipulation with some other kinds of structures, say \( M_1, M_2, \ldots, M_s \).
- Knowledge of all structures of each type \( M_i \) implies in principle the knowledge of all structures \( \gamma \).
- Still special technique for manipulation and isomorphism recognition should be elaborated.
1-factorizations of the graph $K_{2m}$ is a nice example of structures $\gamma$.

Clearly, the size of the complete graph $K_n$ should be even.

For $n = 4$ just one solution.

Can be regarded as coloured graph $\mathcal{M}$.

$Aut(\mathcal{M}) = E_4$, $CAut(\mathcal{M}) = S_4$. 
Example (2.8 One-factorizations of $K_6$)

- We start from scratch.
- The union of any two factors is a graph of valency 2.
- For $n = 6$ there is just the cycle $C_6$:

\[
\begin{array}{c}
\text{\includegraphics[width=0.4\textwidth]{cycle.png}}
\end{array}
\]

- Remove it from $K_6$.
- Get:

\[
\begin{array}{c}
\text{\includegraphics[width=0.4\textwidth]{remaining.png}}
\end{array}
\]
Just one possibility for further split.

Indeed, start from \( \{3, 6\} \) and reject other one-factor, containing it.

Note significant relabelling at the middle of the process.
It is clear that up to isomorphism, there exists just one solution $F$, constructed by us.

It turns out the $\text{Aut}(F) = \text{Id}$, $\text{CAut}(F) \simeq S_5$ (see further lectures).

Therefore the orbit of $\mathcal{M}$ has length 6.

A way to understand two copies of $S_5$ inside of $S_6$:

- intransitive and
- 3-transitive.
Example (2.9)

- One more class of problems where we observe combination of a few different strategies is the enumeration of cubic graphs.
- Has a long history.
- Hereditary effect is significant: "History of search" for small amount of vertices can be used.
- Again, solutions are possible only for even $n$.
- The speed of generation nowadays is surprisingly high.
### Table with history

<table>
<thead>
<tr>
<th>n</th>
<th>Year</th>
<th>Name</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>1889</td>
<td>deVries</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>1966/67</td>
<td>Balaban</td>
<td>85</td>
</tr>
<tr>
<td>14</td>
<td>1976</td>
<td>Bussemaker et al</td>
<td>509</td>
</tr>
<tr>
<td>16</td>
<td>1976</td>
<td>Faradžev</td>
<td>4060</td>
</tr>
<tr>
<td>18</td>
<td>1976</td>
<td>Faradžev</td>
<td>41,301</td>
</tr>
<tr>
<td>20</td>
<td>1986</td>
<td>McKay-Royle</td>
<td>510,489</td>
</tr>
<tr>
<td>22</td>
<td>1992</td>
<td>Brinkmann</td>
<td>7,319,447</td>
</tr>
<tr>
<td>24</td>
<td>1992</td>
<td>Brinkmann</td>
<td>117,940,535</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>Meringer</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>1998</td>
<td>Sanjmyatav-McKay</td>
<td>2,094,480,864</td>
</tr>
</tbody>
</table>
In a more evident form the starting elements of building kit approach in AGT were suggested by Reichard (2003) and Degraer & Coolsaet (2005). Both were interested in imprimitive antipodal distance regular graphs of diameter 3. Starting ingredients are SRGs and spreads in them. This approach will be extended and exploited in lecture 7 by M. Macaj.
Constructive enumeration of association schemes

- The most sophisticated area for enumeration: requirements are described via the tensor of structure constants.
- Compare with, say, one-factorizations and regular graphs.
- Mainly, orderly generation has been used.
- Canonical isomorphs are labelled lexicographically.
- Use of directed graphs is necessary.
In a sense, the simplest case is enumeration of symmetric association schemes (AS) with two classes.
Each class is strongly regular graph (srg).
Usually, one talks about enumeration of srg’s.
Computer aided efforts have been started at Eindhoven and Moscow.
First non-trivial cases appear on 25, 26, and 28 vertices.
Use of two-graphs and Seidel switching is helpful.
Results were obtained around 1975 independently and were agreeing.
The next case \( n = 29 \) was attacked by the late V.A. Zaichenko (Moscow).
He discovered all graphs, though did not prove that the list is complete.
Ted Spence started from $n = 29$ and proved completeness (probably was not aware of Zaichenko’s result).

Full catalogues on up to 40 vertices are obtained and presented on his home page.

Full results also for some cases on 45, 49, 50 and 64 vertices.

Use of two-graphs and other auxiliary structures is essential.

Generation techniques remain in a sense partially hidden.
Each class of a non-symmetric AS with 2 classes is called a doubly regular tournament (DRT).

A student task for Dima Pasechnik was to enumerate everything on 15 points.

He discovered and investigated non-Schurian example (to be mentioned in further lectures).

Two proofs that it is non-Schurian: Via group and via the use of the so-called $t$-condition.
For a while there were no attempts to enumerate all ASs.
This job was done by A. Hanaki and I. Miyamoto.
The results appear on their home page http://math.shinshu-u.ac.jp/~hanaki/as/.
Full enumeration for up to 34 points (except for 31).
Partial results in many other cases.
The main ideology of Hanaki and Miyamoto is described in their papers.

Also some of their programs (to be used in GAP) are available from their home page and proved to be helpful.

Nevertheless, the full panorama of all their computer aided strategy was not described fully and evidently.
- There were just a few attempts to start enumeration of all coherent configurations (CCs).
- Japanese students, Reichard and finally Matan Ziv-Av (subject of Lecture 3).
- It is time for a new wave of healthy competition in this area.
- Hopefully, this summer school will promote efforts in this direction.
Computer package COCO (COherent COnfigurations) was created in 1990-1992 by I.A. Faradžev with theoretical and algorithmical support of MK.

Its main goal is to enumerate all AS that are invariant with respect to a given permutation group \((H, \Omega)\).

Predecessors of COCO were some investigations together with A.A. Ivanov, as well as the Ph.D. thesis of Zaichenko.
Main functions:

- **ind** Inducing of \((H, R)\) on structures of a given type;
- **cgr** Construction of 2-orbits of \((H, \Omega)\) in form of the full colour graph;
- **inm** Construction of the structure constants of \(V(H, \Omega)\);
- **sub** Construction of all mergings of \(V(H, \Omega)\) that are AS;
- **aut** Construction of \(Aut(W)\) for each obtained merging \(W\).
- A few other functions exist that are rarely used.
- During the last 20 years it was used extensively by MK and his colleagues.
- The work on COCO was literally interrupted by the collapse of the USSR.
- The package is available from the home page of A.E. Brouwer.
- Recent small improvements by Dima Pasechnik.
For a while MK was dreaming that a new update COCO-II will appear.

Serious attempts by Christian Pech and Sven Reichard.

COCO-IIP vs COCO-IIR.

Some information on the home page of Pech.

Some attempts of Ziv-Av to “stand in the middle”. 
Main questions are related to ambitions for extension, ideology, practical goals.

Should be able to enumerate all CCs, not only ASs.

More possibilities with Aut, CAut, AAut.

Isomorphism of color graphs.

Ideally a share package of GAP.

“To be or not to be, that is the question.”
Some fulfilled projects

- There are ongoing efforts to reach new computer aided results of constructive enumeration in classical areas of graph theory and design theory.
- Some of them were already mentioned.
- Very striking history for STSs:
  - Classical stage is described by Alex Rosa et al;
  - They exist for $n = 1, 3(\text{mod} 6)$;
  - Kaski-Östergård: $n = 19$;
  - Same team: $n = 21$ with non-trivial group.
• Similar history for regular graphs and one-factorizations.
• A lot of efforts in design theory (again, K&Ö among the leaders).
• Attempts to consider particular cases of SRGs (R. Mathon, C. Lam et al).
• Classical proof of the non-existence of a projective plane of order 10 (Lam et al).
Some projects related to efforts of MK et al:

- enumeration of partial difference sets aka Cayley srgs: Aiso Heinze;
- enumeration of small vertex-transitive directed srgs: F. Fiedler et al;
- new parameter sets for srgs (MK et al);
- subgraphs and equitable partitions in known triangle-free srgs (Ziv-Av);
- some kinds of amorphic association schemes (MK, N. Kriger et al).
During this school we will pay attention to the following projects:

- enumeration of small CCs of order up to 15 (Ziv-Av, Lecture 3);
- enumeration of S-rings over groups of order up to 63 (Ziv-Av, Lecture 5);
- enumeration of so-called Siamese objects, of order 4, 15 and 40 (Macaj, Lecture 7).

Lecture 8 will be devoted to COCO-II (Reichard).
References


DegC05 Degraer, J.; Coolsaet, K. Classification of three-class association schemes using backtracking with dynamic variable ordering. Discrete Math. 300 (2005), no. 1-3, 71–81


HanM http://math.shinshu-u.ac.jp/~hanaki/as/


<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spe</td>
<td><a href="http://www.maths.gla.ac.uk/~es/">http://www.maths.gla.ac.uk/~es/</a></td>
</tr>
</tbody>
</table>
Thank you!