Inhomogeneous hypergraphs

Élie de Panafieu,

Algebra Institute,
J. Kepler University, Linz, Autriche

Seminar of the Complex Networks team
2015
Introduction

Tool: Analytic combinatorics
Model: Inhomogeneous hypergraphs
Application: Constraint Satisfaction Problems (CSP)
Introduction

Tool: Analytic combinatorics

Model: Inhomogeneous hypergraphs

Application: Constraint Satisfaction Problems (CSP)

\[ \exists x_1, \ldots, x_n, \ C_1 (x_{r(1,1)}, \ldots, x_{r(1,a_1)}) \wedge \cdots \wedge C_m (x_{r(m,1)}, \ldots, x_{r(m,a_m)}) \]

Enumeration of satisfied instances.
Introduction

Tool: Analytic combinatorics

Model: Inhomogeneous hypergraphs

Application: Constraint Satisfaction Problems (CSP)

Generating function

\[ A(z) = \sum_{a \in \mathcal{A}} \frac{z^{\mid a \mid}}{\mid a \mid!}, \]

combinatorial description \( \rightarrow \) equation on the generating function,

asymptotics and limit laws \( \leftarrow \) analytic properties.
Inhomogeneous graphs

“The evolution of graphs may be considered as a rather simplified model of the evolution of certain communication nets [...] Of course, if one aims at describing such a real situation, one should replace the hypothesis of equiprobability of all connections by some more realistic hypothesis.”

– Erdős, Rényi (1959)

Probabilistic model introduced by Söderberg, extended by Bollobás, Janson, Riordan. Enumerative model analyzed by E.d.P. and Ravelomanana.

- The model inputs a parameter \( R \in \text{Sym}_q(\mathbb{R}_{\geq 0}) \),
- each vertex \( v \) receives a color \( t(v) \) in \( \{1, \ldots, q\} \),
- each edge \( (v, w) \) has a weight \( R_{t(v), t(w)} \).

Inhomogeneous graphs are counted with a weight

\[
\text{weight}(G) = \prod_{(v, w) \in \text{edges}(G)} R_{t(v), t(w)}.
\]
Inhomogeneous graphs

The model inputs a parameter $R \in \text{Sym}_q(\mathbb{R}_{\geq 0})$, each vertex $v$ receives a color $t(v)$ in $\{1, \ldots, q\}$, each edge $(v, w)$ has a weight $R_{t(v), t(w)}$.

Inhomogeneous graphs are counted with a weight

$$\text{weight}(G) = \prod_{(v, w) \in \text{edges}(G)} R_{t(v), t(w)}.$$
A graph is properly $q$-colored if each vertex has a color in $\{1, \ldots, q\}$, and no edge links two vertices having the same color.

Bijection between inh. graphs and properly $q$-colored graphs

$$ R = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{G = } \quad \text{weight}(G) = 0 $$

Asymptotics for $n$ vertices and $m$ edges

$$ \frac{n^{2m}}{2^m m!} e\left(\frac{m}{n}\right)^2 \frac{q}{q-1} \left(1 + \frac{2}{q - 1} \frac{m}{n}\right)^{-\frac{q-1}{2}} \left(1 - \frac{1}{q}\right)^m q^n \left(1 + o(1)\right) $$

when $m/n$ has a positive limit. A result already obtained by Wright (1972).
Example: properly $q$-colored graphs

A graph is properly $q$-colored if each vertex has a color in $\{1, \ldots, q\}$, and no edge links two vertices having the same color.

Bijection between inh. graphs and properly $q$-colored graphs

$$R = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad G = \begin{array}{c} 1 \\ \hline 1 \\ \hline 1 \end{array} \quad \text{weight}(G) = 1$$

Asymptotics for $n$ vertices and $m$ edges

$$\frac{n^{2m}}{2^mm!} e\left(\frac{m}{n}\right)^2 \frac{q}{q-1} \left(1 + \frac{2}{q-1} \frac{m}{n}\right)^{-\frac{q-1}{2}} \left(1 - \frac{1}{q}\right)^m q^n (1 + o(1))$$

when $m/n$ has a positive limit. A result already obtained by Wright (1972).
A 2-colorable graph with \( c \) components has \( 2^c \) proper colorations.

Structure of inh. graph \( \rightarrow \) enumeration of 2-colorable graphs.

result already obtained by Pittel and Yeum (2004).
Example: friendship graphs

Each vertex has \( r \) hobbies among a set of size \( s \), two vertices can be linked only if they share at least \( t \) hobby.

\[
R = \begin{pmatrix}
 & & & & \\
 & & & & \\
 & & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
\end{pmatrix}
\]

\((r, s, t) = (3, 10, 1)\).

\(R\) is the adjacency matrix of the complement of the Kneser graph.
Example: systems of 2-equations

Consider a finite set $E$ of triplets from $\mathbb{F}_d$, and a system

$$
\exists x_1, \ldots, x_n \in \mathbb{F}_d, \quad \begin{cases}
  a_1 x_{e(1,1)} + b_1 x_{e(1,2)} &= c_1, \\
  \vdots & \\
  a_m x_{e(m,1)} + b_m x_{e(m,2)} &= c_m.
\end{cases}
$$

The probability of satisfiability is expressed using the enumeration of inhomogeneous graphs

$$
R = \begin{pmatrix}
\end{pmatrix}
$$

For 2QXorSAT, $R$ is a Hamming-like matrix.
A **CSP** is a set of Boolean functions taking value in a finite set.

An **instance** is a formula

\[ \exists x_1, \ldots, x_n, \ C_1 (x_{r(1,1)}, \ldots, x_{r(1,a_1)}) \land \cdots \land C_m (x_{r(m,1)}, \ldots, x_{r(m,a_m)}) \]

A **satisfied instance** is a pair instance-solution.

**Bijection between satisfied instances of CSP with clauses of arity 2, and inhomogeneous graphs**

- variable
- constraint
- values
- number of clauses satisfied by \((i, j)\) or \((j, i)\)

vertex,
edge,
colors,
\(R_{i,j}\).
Inhomogeneous hypergraphs

- Each vertex $v$ receives a color $t(v)$ in $\{1, \ldots, q\}$,
- edges can contain more than 2 vertices,
- the weight of an edge depends on the colors it contains.

$\text{orderings}(G) = |\{((5, 2, 3), (4, 1), (2, 3), (1, 5, 3), (1, 6)), \ldots\}|,$

$\text{weight}(G) = \frac{\text{orderings}(G)}{|E(G)|! \prod_{e \in E(G)} |e|!} \prod_{e \in E(G)} \omega_{\tilde{t}(e)}$
Analytic combinatorics of inhomogeneous hypergraphs

Many parameters: for all $\bar{t} \in \mathbb{N}^q$, weight $\omega_{\bar{t}}$ for the edges that contain $t_i$ vertices of type $i$.

$$\Omega(x_1, \ldots, x_q) = \sum_{t_1, \ldots, t_q \geq 0} \omega_{t_1, \ldots, t_q} \frac{x_1^{t_1}}{t_1!} \cdots \frac{x_q^{t_q}}{t_q!}.$$ 

Miracle: statistics of the model reduced to analytic properties of $\Omega$.

Example: tree with root of type $i = \text{root}$ and set of edges with 1 vertex of type $i$ removed and the other replaced by rooted trees.

$$T_i(z) = \sum_n \left( \begin{array}{c} \text{nb trees of size } n \\ \text{root of color } i \end{array} \right) \frac{z^n}{n!}.$$ 

---

![Diagram of a tree with roots and branches labeled with numbers.]
Many parameters: for all \( \bar{t} \in \mathbb{N}^q \), weight \( \omega_{\bar{t}} \) for the edges that contain \( t_i \) vertices of type \( i \).

\[
\Omega(\bar{x}) = \sum_{\bar{t} \in \mathbb{N}^q} \omega_{\bar{t}} \frac{\bar{x}^{\bar{t}}}{\bar{t}!}.
\]

Miracle: statistics of the model reduced to analytic properties of \( \Omega \).

Example: tree with root of type \( i = \text{root} \) and set of edges with 1 vertex of type \( i \) removed and the other replaced by rooted trees.

\[
T_i(z) = \sum_n \left( \text{nb trees of size } n \text{ root of color } i \right) \frac{z^n}{n!}.
\]
Analytic combinatorics of inhomogeneous hypergraphs

Many parameters: for all \( \bar{t} \in \mathbb{N}^q \), weight \( \omega_{\bar{t}} \) for the edges that contain \( t_i \) vertices of type \( i \).

\[
\Omega(\vec{x}) = \sum_{\bar{t} \in \mathbb{N}^q} \omega_{\bar{t}} \frac{\vec{x}^{\bar{t}}}{\bar{t}!}.
\]

Miracle: statistics of the model reduced to analytic properties of \( \Omega \).

Example: tree with root of type \( i = \text{root} \) and set of edges with 1 vertex of type \( i \) removed and the other replaced by rooted trees.

\[
T_i(z) = \sum_n \left( \text{nb trees of size } n \text{ root of color } i \right) \frac{z^n}{n!}.
\]
Many parameters: for all $\bar{t} \in \mathbb{N}^q$, weight $\omega_{\bar{t}}$ for the edges that contain $t_i$ vertices of type $i$.

$$
\Omega(\bar{x}) = \sum_{\bar{t} \in \mathbb{N}^q} \omega_{\bar{t}} \frac{\bar{x}^{\bar{t}}}{\bar{t}!}.
$$

Miracle: statistics of the model reduced to analytic properties of $\Omega$.

Example: tree with root of type $i = \text{root}$ and set of edges with 1 vertex of type $i$ removed and the other replaced by rooted trees.

$$
T_i(z) = ze^{\partial_i \Omega(T_1(z),...,T_q(z))}.
$$
Many parameters: for all $\bar{t} \in \mathbb{N}^q$, weight $\omega_{\bar{t}}$ for the edges that contain $t_i$ vertices of type $i$.

$$\Omega(\bar{x}) = \sum_{\bar{t} \in \mathbb{N}^q} \omega_{\bar{t}} \frac{\bar{x}^{\bar{t}}}{\bar{t}!}.$$  

Miracle: statistics of the model reduced to analytic properties of $\Omega$.

Example: tree with root of type $i = \text{root}$ and set of edges with 1 vertex of type $i$ removed and the other replaced by rooted trees.

$$T_i(z) = ze^{\partial_i \Omega(T(z))}$$
Trees and unicycles

Unrooted trees: \[ U(z) = \sum_i T_i + \Omega(\bar{T}) - \bar{T} \partial \Omega(\bar{T}) \]

Unicycles: \[ V(z) = -\frac{1}{2} \log \det (\text{Id} - \text{diag}(\bar{T})\mathcal{H}_\Omega(\bar{T})) \]

Trees and unicycles: excess \( k = \sum_{e \in \text{edges}} (|e| - 1) - n \)

\[ n! [z^n] \frac{U(z)^{-k}}{(-k)!} e^{V(z)} = \frac{n!}{2i\pi} \int \frac{U(z)^{-k}}{(-k)!} e^{V(z)} \frac{dz}{z^{n+1}} \]

(plot from Flajolet, Sedgewick 2009)
Asymptotics of all inhomogeneous hypergraphs

Sum of the weights of inhygraphs with \( n \) vertices and excess \( k \)

\[
[y^{k+n}] \sum_{n_1 + \cdots + n_q = n} \binom{n}{\bar{n}} \prod_{\bar{t} \in \mathbb{N}^q} \left( 1 + \omega_{\bar{t}} y^{t_1 + \cdots + t_q - 1} \right)^{\binom{n_1}{t_1} \cdots \binom{n_q}{t_q}}
\]

Approximations: coefficient extraction, Stirling, integral for the sum

\[
\sim C_{n,k} \int_{\sum x_i = 1; x_i \in [0,1]^q} A(\bar{x}) e^{n\Phi(\bar{x})} d\bar{x}
\]

where \( \Psi(\bar{x}) = \bar{x} \partial_{\bar{x}} \Omega(\bar{x}) - \Omega(\bar{x}) \sum_{i=1}^q x_i \), \( \Psi(\bar{\zeta}) = 1 + \frac{k}{n} \)

and \( \Phi(\bar{x}) = \bar{x} (\log(\bar{\nu}) - \log(\bar{x})) + \frac{\Omega(\bar{\zeta})}{\zeta} - (1 + \frac{k}{n}) \log(\zeta) \).

<table>
<thead>
<tr>
<th>trees and unicycles</th>
<th>bounded excess components</th>
<th>giant component</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0 )</td>
<td>( \Lambda + O(n^{-1/3}) )</td>
<td>( \sum_{e \in \text{edges}} (</td>
</tr>
</tbody>
</table>
Asymptotics of all inhomogeneous hypergraphs

Sum of the weights of inhygraphs with \( n \) vertices and excess \( k \)

\[
[y^{k+n}] \sum_{n_1 + \cdots + n_q = n} \left( \frac{n}{\bar{n}} \right) \exp \left( \frac{\Omega(y\bar{n})}{y} \right)
\]

Approximations: coefficient extraction, Stirling, integral for the sum

\[
\sim C_{n,k} \int_{x_1 + \cdots + x_q = 1} A(\bar{x}) e^{\Phi(\bar{x})} d\bar{x}
\]

where \( \Psi(\bar{x}) = \frac{\bar{x} \sum_{i=1}^{q} x_i - \Omega(\bar{x})}{\sum_{i=1}^{q} x_i}, \quad \Psi(\zeta \bar{x}) = 1 + \frac{k}{n} \)

and \( \Phi(\bar{x}) = \bar{x} (\log(\nu) - \log(\bar{x})) + \frac{\Omega(\zeta \bar{x})}{\zeta} - (1 + \frac{k}{n}) \log(\zeta) \).

trees and unicycles bounded excess components giant component


<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0 )</td>
<td>( \Lambda + O(n^{-1/3}) )</td>
<td>( \Gamma_1 \cdots \Gamma_q )</td>
<td>( \sum_{e \in \text{edges}} (</td>
</tr>
</tbody>
</table>


Asymptotics of all inhomogeneous hypergraphs

Sum of the weights of inhygraphs with \( n \) vertices and excess \( k \)

\[
[y^{k+n}] \sum_{n_1 + \cdots + n_q = n} \binom{n}{\bar{n}} \exp \left( \frac{\Omega(y\bar{n})}{y} \right)
\]

Approximations: coefficient extraction, Stirling, integral for the sum

\[
\sim C_{n,k} \int_{\bar{x} \in [0,1]^q} A(\bar{x}) e^{n\Phi(\bar{x})} d\bar{x}
\]

where \( \Psi(\bar{x}) = \frac{\bar{x}\tilde{\delta}\Omega(\bar{x}) - \Omega(\bar{x})}{\sum_{i=1}^q x_i} \), \( \Psi(\zeta \bar{x}) = 1 + \frac{k}{n} \)

and \( \Phi(\bar{x}) = \bar{x} \left( \log(\nu) - \log(\bar{x}) \right) + \frac{\Omega(\zeta \bar{x})}{\zeta} - \left( 1 + \frac{k}{n} \right) \log(\zeta) \).
Asymptotics of all inhomogeneous hypergraphs

Sum of the weights of inhygraphs with \( n \) vertices and excess \( k \)

\[ [y^{k+n}] \sum_{n_1 + \cdots + n_q = n} \binom{n}{\bar{n}} \exp \left( \frac{\Omega(y\bar{n})}{y} \right) \]

Approximations: coefficient extraction, Stirling, integral for the sum

\[ \sim C_{n,k} \int_{\bar{x} \in [0,1]^q} A(\bar{x}) e^{n\Phi(\bar{x})} d\bar{x} \]

where

\[ \Psi(\bar{x}) = \frac{\bar{x} \delta \Omega(\bar{x}) - \Omega(\bar{x})}{\sum_{i=1}^{q} x_i}, \quad \Psi(\zeta\bar{x}) = 1 + \frac{k}{n} \]

and

\[ \Phi(\bar{x}) = \bar{x} (\log(\nu) - \log(\bar{x})) + \frac{\Omega(\zeta\bar{x})}{\zeta} - (1 + \frac{k}{n}) \log(\zeta). \]
Asymptotics of all inhomogeneous hypergraphs

Sum of the weights of inhygraphs with $n$ vertices and excess $k$

$$[y^{k+n}] \sum_{n_1+\cdots+n_q=n} \binom{n}{\bar{n}} \exp \left( \frac{\Omega(y\bar{n})}{y} \right)$$

Approximations: coefficient extraction, Stirling, integral for the sum

$$\sim C_{n,k} \int_{x_1+\cdots+x_q=1} \frac{A(\bar{x})e^{n\Phi(\bar{x})}}{\bar{x}^{\prod_{i=1}^{q} x_i}} \, d\bar{x}$$

where

$$\Psi(\bar{x}) = \frac{\bar{x} \bar{\partial} \Omega(\bar{x}) - \Omega(\bar{x})}{\sum_{i=1}^{q} x_i}, \quad \Psi(\zeta \bar{x}) = 1 + \frac{k}{n}$$

and

$$\Phi(\bar{x}) = \bar{x} \left( \log(\bar{\nu}) - \log(\bar{x}) \right) + \frac{\Omega(\zeta \bar{x})}{\zeta} - (1 + \frac{k}{n}) \log(\zeta).$$

trees and unicycles | bounded excess components | giant component
-------------------|-----------------------------|---------------------
0 | $\Lambda + O(n^{-1/3})$ | $\Gamma_1 \cdots \Gamma_q \frac{\sum_{e \in \text{edges}} (|e| - 1)}{n}$
Applications will provide new inhygraphs properties to investigate

- structure of graphs with a giant component
generating function of connected graphs with the same density of edges,

- graphs with forbidden subgraphs
  satisfiability threshold of 2-SAT,

- data base modelling using inhomogeneous hypergraphs
  analysis of data mining algorithms (Peps HYDrATA).