Generating functions of graphs, directed graphs and 2-SAT

Élie de Panafieu [a] *Sergey Dovgal* [b] Dimbinaina Ralaivaosaona [c] Vonjy Rasendrahasina [d] Vlady Ravelomanana [e] Alexandros Singh [f] Stephan Wagner [g]

[a] Bell Labs, Nokia

[b] LaBRI, Université de Bordeaux

[c] Stellenbosch University

[d] Antananarivo, Madagascar

[e] IRIF, Université de Paris

[f] LIPN, Université Sorbonne Paris Nord

[g] University of Uppsala

Seminaire Modèles et Algorithmes, LIGM, 30/03/2021 (online)

Part I. Generating functions and simple graphs

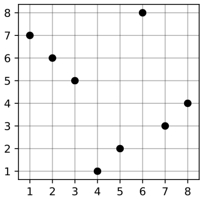
Generating functions and the symbolic method

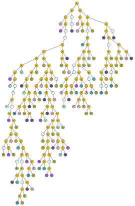
Framework

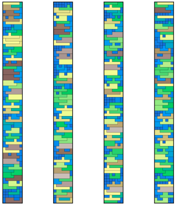
- Graphs are discrete objects with labeled vertices and edges
- The *size* of the graph is the number of its vertices
- The *secondary parameter* is the number of edges
- Let $a_{n,m}$ be the number of graphs with n vertices and m edges

Generating function of the counting sequence:

$$A(z, w) = \sum_{n,m=0}^{\infty} a_{n,m} \frac{z^n}{n!} w^m$$

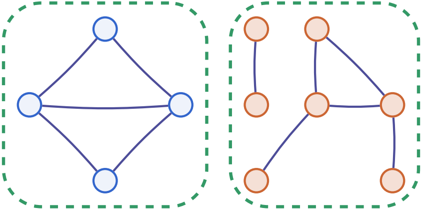






The cartesian product

$$\left(a_0 + a_1 \frac{z}{1!} + a_2 \frac{z^2}{2!} + \dots\right)\left(b_0 + b_1 \frac{z}{1!} + b_2 \frac{z^2}{2!} + \dots\right) = c_0 + c_1 \frac{z}{1!} + c_2 \frac{z^2}{2!} + \dots$$



The convolution rule corresponding to EGF:

$$c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}$$

The Sets, Sequences and Cycles

- $A(z) \times A(z)$ is a *ordered pair* of generating functions
- A *sequence* (ordered)

$$\mathrm{Seq}_A(z) = 1 + A(z) + A(z)^2 + A(z)^3 + \ldots = \frac{1}{1 - A(z)}$$

- $\frac{1}{2}A(z)^2$ is an *unordered pair*
- A *set* (non-ordered)

$$\mathrm{Set}_A(z) = 1 + A(z) + \frac{A(z)^2}{2!} + \frac{A(z)^3}{3!} + \ldots = e^{A(z)}$$

- A *cycle* is a sequence modulo cyclic shifts

$$\mathrm{Cyc}_A(z) = A(z) + \frac{A(z)^2}{2} + \frac{A(z)^3}{3} + \ldots = \log \frac{1}{1 - A(z)}$$

Pointed derivative

Theorem

Let $A(z) = \sum_{n \geqslant 0} a_n \frac{z^n}{n!}$ *count a family* $\mathcal{A}$ *of graphs . Let*

$$A^{\bullet}(z) := z\partial_z A(z) = \sum_{n \geqslant 0} n a_n \frac{z^n}{n!}.$$

Then, $A(z)$ *counts graphs from* $\mathcal{A}$ *with one distinguished vertex.*

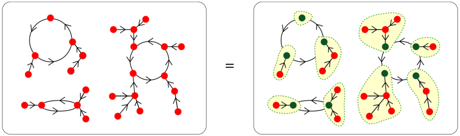
Cayley formula for trees

Theorem

The number of labelled trees with n vertices is n^{n-2}.

Proof. Let $T(z)$ and $U(z)$ count rooted and unrooted trees.

- The number of endomorphisms $[n] \to [n]$ is n^n.
- Let $E(z) = \frac{n^n}{n!}$ be the EGF of endomorphisms.
- Then, $E(z) = \text{Set}(\text{Cyc}(T(z)))$, therefore,

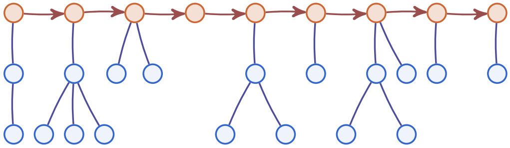


$$E(z) = \exp \log \frac{1}{1 - T(z)} = \frac{1}{1 - T(z)}$$

Cayley formula for trees (continued)

- Also trees:

$$U^{\bullet\bullet}(z) = \mathrm{Seq}_T(z) = \frac{1}{1 - T(z)}$$



- Therefore,

$$\sum_{n=0}^{\infty} n^2 u_n \frac{z^n}{n!} = \frac{1}{1 - T(z)} = \sum_{n=0}^{\infty} n^n \frac{z^n}{n!}.$$

- The number of *rooted* labelled trees is n^{n-1}.

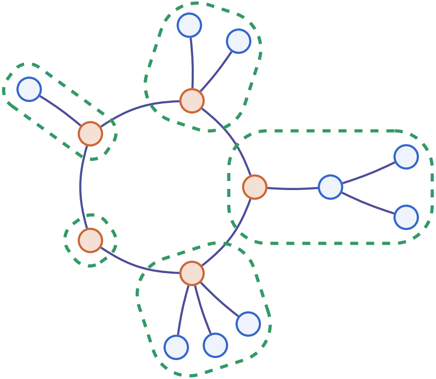
Rooted, Unrooted and Monocyclic

Theorem

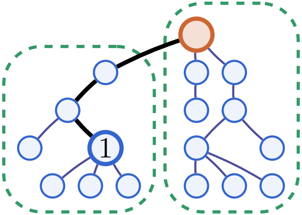
The EGF of Rooted and Unrooted trees, and Monocycles $T(z)$, $U(z)$ and $M(z)$ satisfy

$$
\begin{aligned}
T(z) &= ze^{T(z)} \\
U(z) &= T(z) - \frac{T(z)^2}{2} \\
M(z) &= \frac{1}{2}\left[\log\frac{1}{1-T(z)} - T(z) - \frac{T(z)^2}{2}\right]
\end{aligned}
$$

Unrooted trees and Monocycles (continued)

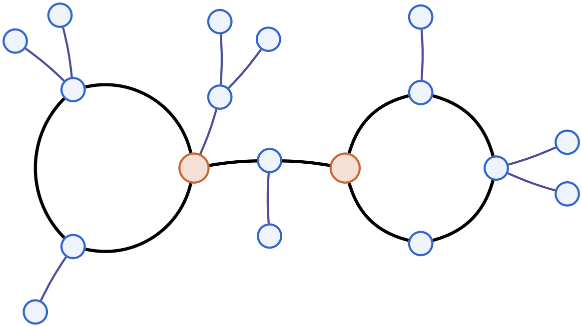


$$M(z) = \mathrm{Cyc}_{\geqslant 3}(T(z))$$



$$T(z) = U(z) + \frac{T(z)^2}{2}$$

Bicycles

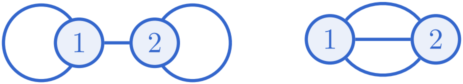


$$B(z) = \frac{1}{8}\frac{1}{(1 - T(z))^3} - \varepsilon(z)$$

The function $\varepsilon(z)$ includes “corner cases”, where the length of one of the cycles is less than 3.

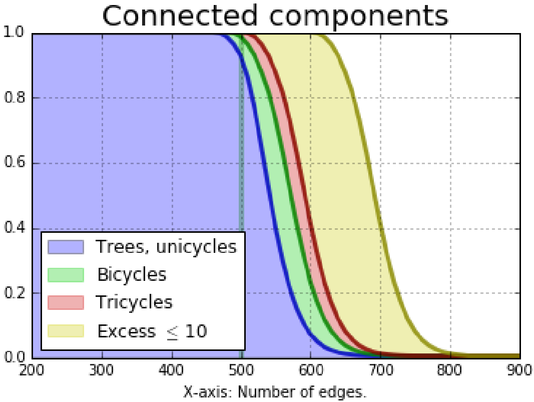
Bicycles (continued)

There are essentially two types of bicycles (a finite number of types)



$$
\begin{aligned}
B(z) &= B_{\bigcirc 1 - 2 \bigcirc}(z) + B_{1 \ominus 2}(z) \\
&= \left(\frac{1}{8} + \frac{1}{12}\right) \frac{1}{(1 - T(z))^3} - \varepsilon(z) \\
&= \frac{5}{24} \frac{1}{(1 - T(z))^3} - \varepsilon(z)
\end{aligned}
$$

Phase transition



Symbolic constructions

- Unrooted Trees and Monocycles $e^{U(z)+M(z)}$
- With one additional bicycle $e^{U(z)+M(z)}B(z)$

Part II. Above the threshold

How about many-cycles?

Δ-multigraphs

Let $\mathcal{D}$ be the set of allowed vertex degrees, and $\Delta(x) = \sum_{d \in \mathcal{D}} \frac{x^d}{d!}$.

Proposition

The number of Δ-multigraphs with n vertices and m multiedges is

$$\mathrm{MG}_{\Delta,n,m} = (2m)![x^{2m}]\Delta(x)^n.$$

Proof. A multigraph is obtained by linking its half-edges.

Excess counting

Definition

A multigraph with n vertices and $n+k$ edges has an excess k. (Bicycles have excess 1, Tricycles have excess 2, etc.)

Proposition

The generating function of Δ-multigraphs (not necessarily connected) of excess k is equal to

$$\mathrm{MG}_{\Delta,k}(z) = \frac{2^k}{\sqrt{\pi}}\Gamma(k+1/2)[x^{2k}]\frac{1}{\left(1-2zx^{-2}\Delta(x)\right)^{k+1/2}},$$

where

$$\frac{2^k}{\sqrt{\pi}}\Gamma(k+1/2) = \begin{cases} \frac{(2k)!}{2^k k!} & \text{if } k \geqslant 0, \\ (-1)^k \frac{2^{|k|}|k|!}{(2|k|)!} & \text{otherwise.} \end{cases}$$

Proof. Coefficient manipulation and magic.

Excess counting (continued)

Proposition

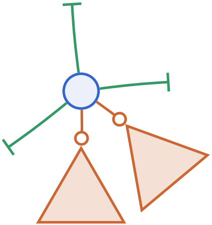
The EGF of Δ-multigraphs of excess k with an additional variable u marking the connected components that are trees is

$$\mathrm{MG}_{\Delta,k}(z,u) = \frac{2^k}{\sqrt{\pi}}\Gamma(k+1/2)[x^{2k}]\frac{1}{\left(1-2x^{-2}(f(z,x)+uU(z))\right)^{k+1/2}},$$

where

$$f(z,x) = z\Delta(x+T(z)) - z\Delta(T(z)) - xT(z).$$

Proof. $f(z,x)$ is EGF of one vertex with attached half-edges and trees



+ Coefficient manipulation and magic

Marking vertices in the complex component

Proposition

The generating function of Δ-multigraphs of excess k with additional variables u, v, z_u, z_v, z_{cx} marking respectively the numbers of trees, unicycles, vertices in trees, in unicycles, and in the complex part, is
$\mathrm{MG}_{\Delta,k}(z, u, v, z_u, z_v, z_{cx}) =$

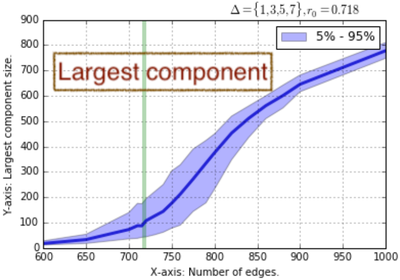
$$\frac{2^k}{\sqrt{\pi}}\Gamma(k+1/2)[x^{2k}]\frac{e^{-M(z_{cx}z)+vM(z_v z)}}{\left(1-2x^{-2}(f(z_{cx}z,x)+uU(z_u z))\right)^{k+1/2}},$$

where

$$f(z,x) = z\Delta(x+T(z)) - z\Delta(T(z)) - xT(z)$$

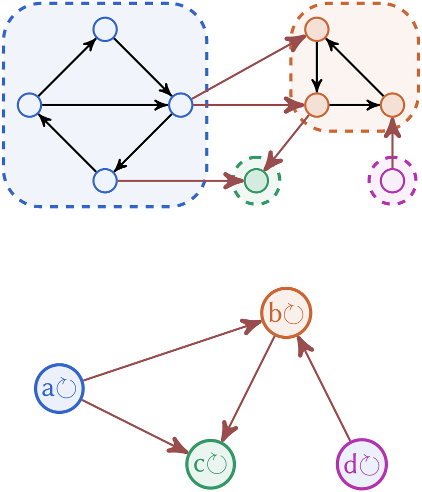
Extensive parameter list available for marking

- Number of trees
- Vertices in trees
- Number of unicycles
- Vertices in unicycles
- Vertices in the “complex” part
- The *core* and the *kernel* (pruning the trees and smoothing the paths)
- * The number of connected multi-cyclic components (if small) by inclusion-exclusion: direct access to the *giant component*



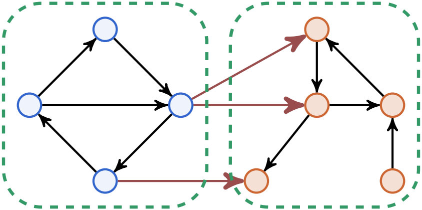
Part III. Strongly connected components in directed graphs

Directed graphs and their components



- Components a↻ , b↻ , c↻ and d↻ are *strongly-connected* components.
- Components a↻ and d↻ are *source-like* components
- Component c↻ is a *sink-like* component

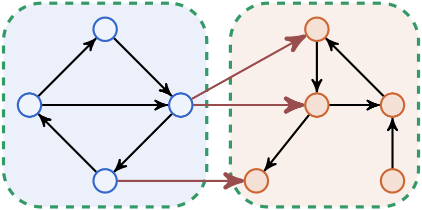
The arrow product



The arrow product convolution rule

Let $a_n = a_n(w)$, $b_n = b_n(w)$, $c_n = c_n(w)$,

$$\left(\sum_{n=0}^{\infty} a_n \frac{1}{(1+w)^{\binom{n}{2}}} \frac{z^n}{n!} \right) \left(\sum_{n=0}^{\infty} b_n \frac{1}{(1+w)^{\binom{n}{2}}} \frac{z^n}{n!} \right) = \sum_{n=0}^{\infty} c_n \frac{1}{(1+w)^{\binom{n}{2}}} \frac{z^n}{n!}$$



The convolution rule corresponding to Graphic GF:

$$c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k} (1+w)^{k(n-k)}$$

Conversion between Exponential GF and Graphic GF

$$\widehat{A}(z,w) = \sum_{n=0}^{\infty} a_n(w) \frac{1}{(1+w)^{\binom{n}{2}}} \frac{z^n}{n!}, \quad A(z,w) = \sum_{n=0}^{\infty} a_n(w) \frac{z^n}{n!}$$

- **Exponential Hadamard product:**

$$\left(\sum_{n\geqslant 0} a_n(\mathrm{w}) \frac{z^n}{n!}\right) \odot_z \left(\sum_{n\geqslant 0} b_n(\mathrm{w}) \frac{z^n}{n!}\right) := \sum_{n\geqslant 0} a_n(\mathrm{w}) b_n(\mathrm{w}) \frac{z^n}{n!}$$

- **Exponential GF for graphs, Graphic GF for sets:**

$$G(z,w) = \sum_{n\geqslant 0} (1+w)^{\binom{n}{2}} \frac{z^n}{n!}, \quad \widehat{Set}(z,w) = \sum_{n\geqslant 0} \frac{1}{(1+w)^{\binom{n}{2}}} \frac{z^n}{n!},$$

- **Conversion formulas:**

$$\boxed{A(z,w) = G(z,w) \odot_z \widehat{A}(z,w)} \quad \boxed{\widehat{A}(z,w) = \widehat{Set}(z,w) \odot_z A(z,w)}$$

Main enumeration theorem

[Liskovets, Robinson, Gessel, Wright et. al. '1970's] [de Panafieu, D. '2019]

Theorem (rediscovery of the results from '1970s)

- *Graphic GF for digraphs with strongly connected components from given family $\mathcal{SCC}$ is*

$$\widehat{D}(z, w) = \frac{1}{e^{-\mathrm{SCC}(z,w)} \odot_z \widehat{Set}(z, w)}$$

where $\mathrm{SCC}(z, w)$ *is the Exponential GF.*

Compare with simple graphs (folklore)

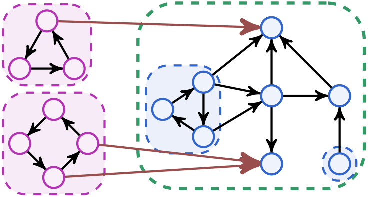
- Exponential GF for graphs with connected components from given family $\mathcal{C}$ is

$$G(z, w) = e^{C(z,w)} = \frac{1}{e^{-C(z,w)}}$$

Proof of the main enumeration theorem

- Let u mark *source-like components* in $\mathcal{D}$.
- $\mathcal{D}$ with *distinguished* source-like components is an **arrow product** of a *set* of strong components and $\mathcal{D}$.

$$\widehat{D}(z, w, \mathrm{u} + \mathbf{1}) = \left(e^{\mathrm{u}\cdot\mathbf{SCC}(z,w)} \odot_z \widehat{Set}(z, w)\right) \cdot \widehat{D}(z, w, 1).$$



are distinguished source-like

are other source-like components.

- Set u $= -1$. Result follows: $\widehat{D}(z, w) = \dfrac{1}{e^{-\mathbf{SCC}(z,w)} \odot_z \widehat{Set}(z, w)}$.

Corollary: strongly connected digraphs

[Liskovets, Robinson, Gessel, Wright et. al. '1970's] [de Panafieu, D. '2019]

Theorem

Exponential GF of strongly connected digraphs is

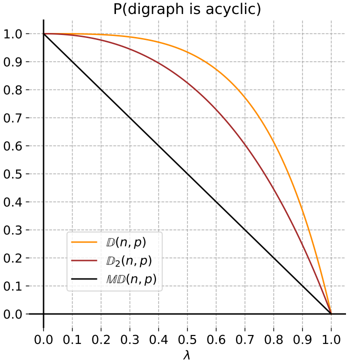
$$\mathrm{SCC}(z, w) = -\log\left(G(z, w) \odot_z \frac{1}{G(z, w)} \right)$$

Proof. Inversion of the main enumeration theorem

$$G(z, w) = \widehat{D}(z, w) = \frac{1}{e^{-\mathrm{SCC}(z,w)} \odot_z \widehat{Set}(z, w)}$$

Graphic GF of all digraphs $\widehat{D}(z, w)$ equals the EGF of graphs $G(z, w)$.

Directed acyclic graphs



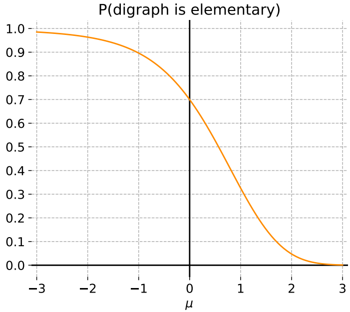
DAGs

A digraph is *directed acyclic* if its connected components are only single vertices

$$\mathrm{SCC}(z, w) = z$$

$$\mathrm{DAG}(z, w) = \frac{1}{\sum_{n \geqslant 0} (1 + w)^{-\binom{n}{2}} \frac{z^n}{n!}}$$

Phase transition in directed graphs



Elementary digraphs

A digraph is called *elementary* if its connected components are only single vertices or cycles

$$\mathrm{SCC}(z, w) = z + \ln \frac{1}{1 - zw} - \varepsilon(z, w)$$

$$\mathrm{Elem}(z, w) = \frac{1}{e^{-\mathrm{SCC}(z,w)} \odot_z \widehat{Set}(z, w)}$$

Part IV. Satisfiable 2-CNF

Exhibition result

Theorem. Let $S_{n,m}$ be the number of satisfiable 2-CNF with n Boolean literals and m clauses. Then,

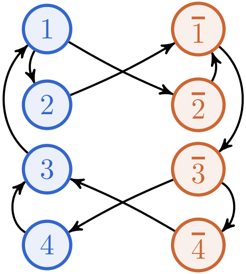
$$\ddot{S}(z,w) = \left[\sqrt{G(z,w) \odot_z \frac{1}{G(2z,w)}} \odot_z \ddot{Set}(z,w) \right] G\left(\frac{2z}{1+w}, w\right)$$

where

- $\ddot{S}(z,w) := \sum_{n=0}^{\infty} \sum_{m=0}^{2n(n-1)} S_{n,m} \frac{w^m}{(1+w)^{n^2}} \frac{z^n}{n!}$
- $\ddot{Set}(z,w) := \sum_{n=0}^{\infty} \frac{1}{(1+w)^{n^2}} \frac{z^n}{n!}$
- $G(z,w) := \sum_{n=0}^{\infty} (1+w)^{\binom{n}{2}} \frac{z^n}{n!}$ is the EGF of all simple graphs
- $\odot_z$ is the exponential Hadamard product $\left(\sum_{n=0}^{\infty} a_n(\mathrm{w}) \frac{z^n}{n!}\right) \odot_z \left(\sum_{n=0}^{\infty} b_n(\mathrm{w}) \frac{z^n}{n!}\right) := \sum_{n=0}^{\infty} a_n(\mathrm{w}) b_n(\mathrm{w}) \frac{z^n}{n!}$.

Implication digraphs

$$\begin{cases} \neg x_1 \vee \neg x_2 & = 1 \\ \neg x_1 \vee x_2 & = 1 \\ x_3 \vee x_4 & = 1 \\ x_1 \vee \neg x_3 & = 1 \\ x_3 \vee \neg x_4 & = 1 \end{cases}$$



Replace each clause $x \vee y$ with two implications $\bar{x} \to y$ and $\bar{y} \to x$.

Proposition (folklore / [Aspvall, Plass, Tarjan '82])

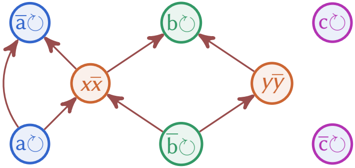
2-CNF is satisfiable if and only if there is no *contradictory circuit*.

The above 2-CNF is not satisfiable

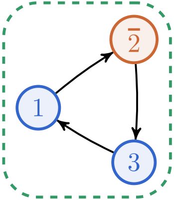
$$1 \longrightarrow \bar{2} \longrightarrow \bar{1} \longrightarrow \bar{3} \longrightarrow \bar{4} \longrightarrow 3 \longrightarrow 1$$

N.B. Each variable of a contradictory component belongs to a contradictory circuit.

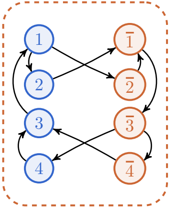
Implication digraphs and their components



2-CNF implication digraph



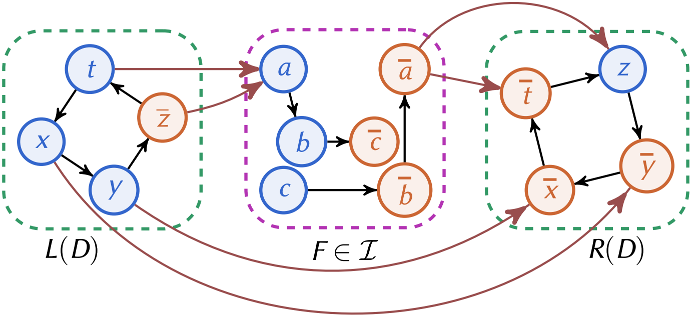
(Ordinary)



(Contradictory)

- $x\bar{x}$ and $y\bar{y}$ are *contradictory* strongly connected
- a and $\bar{b}$ are *ordinary* source-like
- $\bar{a}$ and b are *ordinary* sink-like
- c and $\bar{c}$ are *ordinary isolated*

The implication product

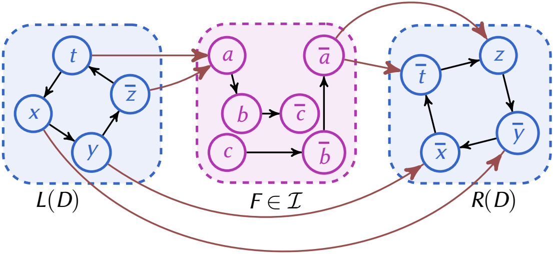


The implication product convolution rule

If $\widehat{A}$ is Graphic GF and $\ddot{B}, \ddot{C}$ are Implication GF then

$$\Big(\sum_{n=0}^{\infty} a_n \frac{2^n}{(1+w)^{\binom{n+1}{2}}} \frac{z^n}{n!}\Big)\Big(\sum_{n=0}^{\infty} b_n \frac{1}{(1+w)^{n^2}} \frac{z^n}{n!}\Big) = \sum_{n=0}^{\infty} c_n \frac{1}{(1+w)^{n^2}} \frac{z^n}{n!}$$

$$\widehat{A}\Big(\frac{2z}{1+w}\Big) \cdot \ddot{B}(z,w) = \ddot{C}(z,w)$$



Combinatorial convolution rule corresponding to Implication GF:

$$c_n = \sum_{k=0}^{n} \binom{n}{k} 2^k a_k b_{n-k} (1+w)^{k \cdot 2(n-k) + \binom{k}{2}}$$

Main enumeration theorem for 2-CNF

Theorem ([de Panafieu, D., Ravelomanana '2021])

- *Implication GF for implication digraphs with ordinary components from given $\mathcal{SCC}$ and contradictory components from given $\mathcal{CSCC}$ is*

$$\ddot{CNF}_2(z,w) = \frac{e^{\mathrm{CSCC}(z,w) - \mathrm{SCC}(2z,w)/2} \odot_z \ddot{Set}(z,w)}{e^{-\mathrm{SCC}\left(\frac{2z}{1+w},w\right)} \odot_z \widehat{Set}(z,w)}$$

- *where*

$$\widehat{Set}(z,w) = \sum_{n\geqslant 0} \frac{1}{(1+w)^{\binom{n}{2}}} \frac{z^n}{n!}, \quad \ddot{Set}(z,w) = \sum_{n\geqslant 0} \frac{1}{(1+w)^{n^2}} \frac{z^n}{n!}.$$

Proof of the main enumeration theorem

- Let u mark *ordinary source-like components* in 2-CNF.
- Let v mark *ordinary isolated components* in 2-CNF (by pairs).
- Take **implication product** of *set* of ordinary components and 2-CNF.
- Add an arbitrary subset of ordinary isolated components. EGF of one pair of isolated components is $\mathrm{SCC}(2z)/2$.
- Now, every source-like component is marked by u or 1, and every ordinary isolated pair is marked by 2u, v or 1.

$$egf[\ddot{\mathrm{CNF}}_2(z, \mathrm{u}+1, 2\mathrm{u}+\mathrm{v}+1)] \\ = egf\Big[\Big(e^{\mathrm{u}\cdot \mathrm{SCC}\left(\frac{2z}{w+1}\right)} \odot_z \widehat{Set}(z)\Big)\ddot{\mathrm{CNF}}_2(z,1,1)\Big] \cdot e^{\mathrm{v}\mathrm{SCC}(2z)/2}$$

- Let $\mathrm{u} = -1$. **An implication digraph without source-like ordinary components is a disjoint set of contradictory and ordinary components.**
- Complete with arithmetic transformations.

Two corollaries

Corollary 1 (inversion of the main theorem)

Exponential GF of contradictory strongly connected components is

$$\mathrm{CSCC}(z, w) = \frac{1}{2}\mathrm{SCC}(2z, w) + \log\left(\mathrm{BG}(z, w) \odot_z \left[\frac{\ddot{\mathrm{CNF}}_2(z, w)}{G\left(\frac{2z}{w+1}\right)}\right]\right)$$

where $\mathrm{BG}(z) = \sum_{n=0}^{\infty}(1 + w)^{n^2}\dfrac{z^n}{n!}$ is the EGF of bipartite graphs.

Corollary 2 (no contradictory components)

Implication GF of satisfiable 2-CNF is

$$\ddot{S}(z, w) = \left[\sqrt{G(z, w) \odot_z \frac{1}{G(2z, w)}} \odot_z \ddot{Set}(z, w)\right] G\left(\frac{2z}{1+w}, w\right).$$

Summary

Discussed today

1. Generating functions of
   - Graphs constructed below the threshold of the phase transition
   - (Multi-)graphs above the threshold (giant component)
   - Directed graphs (with given strong components)
   - 2-Conjunctive Normal Forms
2. Symbolic method for 2-CNF is 2 month old; satisfiable formulas is a particular case

Thank you for your attention.