

# Pointing, asymptotics, and random generation in unlabelled classes

Éric Fusy

*LIX, École Polytechnique (Paris)*

**Joint work** with Manuel Bodirsky, Mihyun Kang  
and Stefan Vigerske

# Motivations

Automatic methods for

- Enumeration (exact/asymptotic)
- Random generation (cf [Flajolet,F,Pivoteau'07])

in the unlabelled setting.

# Motivations

Automatic methods for

- Enumeration (exact/asymptotic)
- Random generation (cf [Flajolet,F,Pivoteau'07])

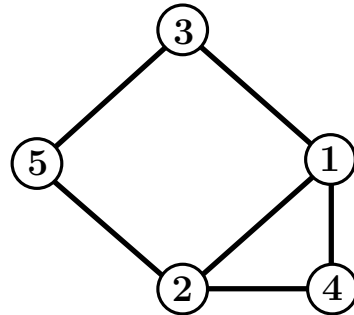
in the unlabelled setting.

## References:

- Short version in SODA'07
- Long version written in the framework of “Combinatorial species”, cf [Bergeron, Labelle, Leroux'98]

# Labelled/Unlabelled structures

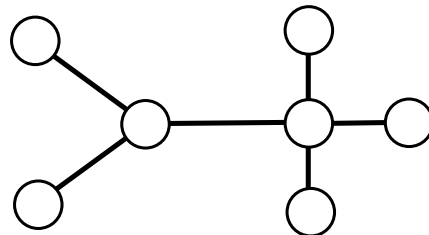
- **labelled** class  $\mathcal{C} = \cup_n \mathcal{C}_n$



**Labeled** graph of size 5

$$\text{EGF : } C(x) = \sum_n \frac{1}{n!} c_n x^n, \text{ with } c_n = |\mathcal{C}_n|$$

- **Unlabelled** class  $\tilde{\mathcal{C}} = \cup_n \tilde{\mathcal{C}}_n$



**Unlabeled** tree of size 7

$$\text{OGF : } C(x) = \sum_n \tilde{c}_n x^n, \text{ with } \tilde{c}_n = |\tilde{\mathcal{C}}_n|$$

# Plan

- Decomposition strategy for labelled structures
  - Pointing + recursive decomp. + gen. functions
  - Examples: trees, planar graphs...

# Plan

- **Decomposition** strategy for **labelled** structures
  - **Pointing** + recursive decomp. + gen. functions
  - Examples: trees, planar graphs...
- We adapt the method to the **unlabelled** setting
  - Difficulties due to **symmetries**
  - Solution: **unbiased pointing** + Pólya theory

# Plan

- **Decomposition** strategy for **labelled** structures
  - **Pointing** + recursive decomp. + gen. functions
  - Examples: trees, planar graphs...
- We adapt the method to the **unlabelled** setting
  - Difficulties due to **symmetries**
  - Solution: **unbiased pointing** + Pólya theory
- Application to **asymptotic enumeration**

# Plan

- **Decomposition** strategy for **labelled** structures
  - **Pointing** + recursive decomp. + gen. functions
  - Examples: trees, planar graphs...
- We adapt the method to the **unlabelled** setting
  - Difficulties due to **symmetries**
  - Solution: **unbiased pointing** + Pólya theory
- Application to **asymptotic enumeration**
- Application to **random generation**:  
⇒ **Boltzmann samplers without rejection**



# Decomposition strategy for labelled structures

# Dictionary for EGF

- labelled class  $\mathcal{C} = \cup_n \mathcal{C}_n$

$$\text{EGF : } C(x) = \sum_n \frac{1}{n!} c_n x^n, \text{ with } c_n = |\mathcal{C}_n|$$

- Simple **computation rule** for each **construction**:

Disjoint union	$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$C(x) = A(x) + B(x)$
Cartesian product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(x) = A(x) \cdot B(x)$
Set	$\mathcal{C} = \text{Set}(\mathcal{A})$	$C(x) = \exp(A(x))$
Substitution	$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$	$C(x) = A(B(x))$

# Dictionary for EGF

- labelled class  $\mathcal{C} = \cup_n \mathcal{C}_n$

$$\text{EGF : } C(x) = \sum_n \frac{1}{n!} c_n x^n, \text{ with } c_n = |\mathcal{C}_n|$$

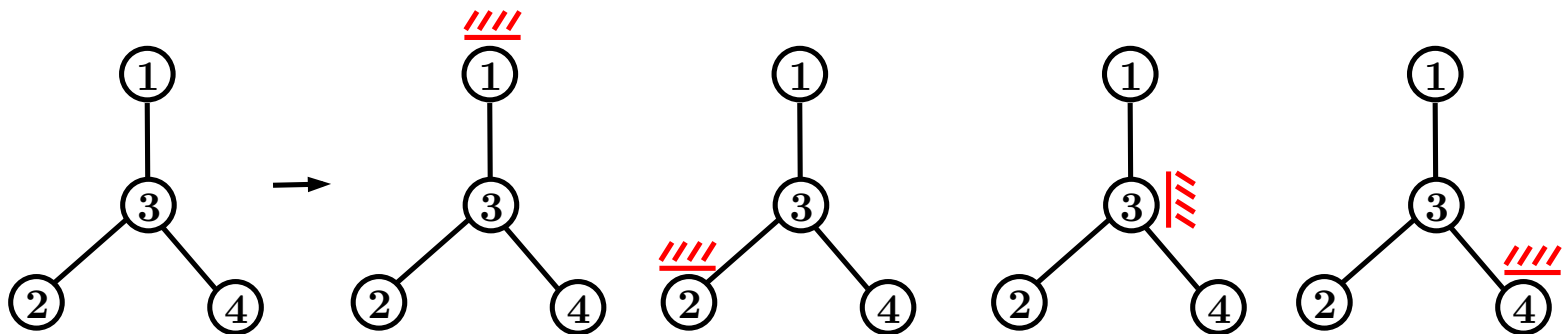
- Simple **computation rule** for each **construction**:

Disjoint union	$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$C(x) = A(x) + B(x)$
Cartesian product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(x) = A(x) \cdot B(x)$
Set	$\mathcal{C} = \text{Set}(\mathcal{A})$	$C(x) = \exp(A(x))$
Substitution	$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$	$C(x) = A(B(x))$

- Remark.** **Substitution** rule implies **Set** rule since the EGF of the class Set is  $\exp(z)$  (same for cycle, set, unoriented sequence, etc...)

# Decomposition strategy for trees

- Goal: find  $t_n$  the number of (unrooted) trees of size  $n$
- Important tool: pointing:  $\mathcal{A} \mapsto \mathcal{A}^\bullet$   
Let  $r_n$  be the number of rooted trees of size  $n$

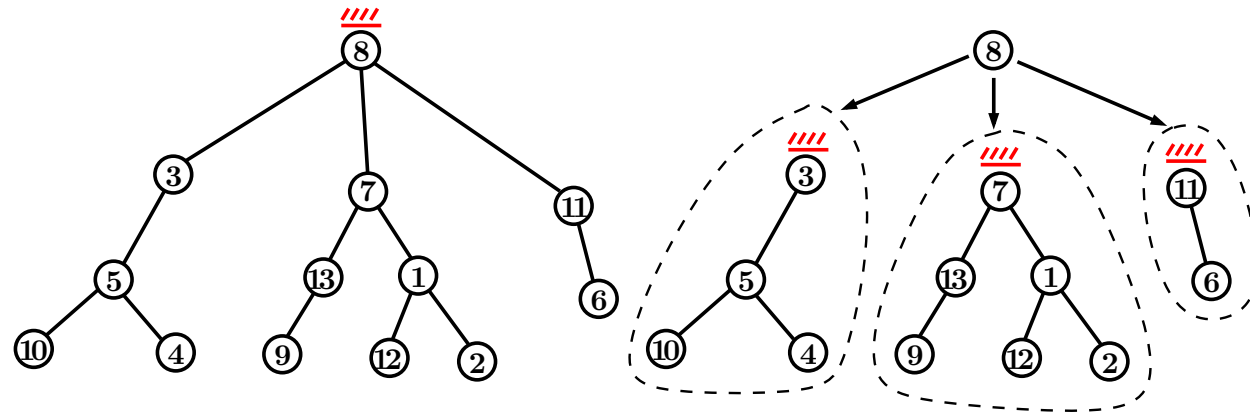


$$r_n = n \cdot t_n$$

$\Rightarrow$  Counting trees reduces to counting rooted trees.

# Rooted trees are decomposable

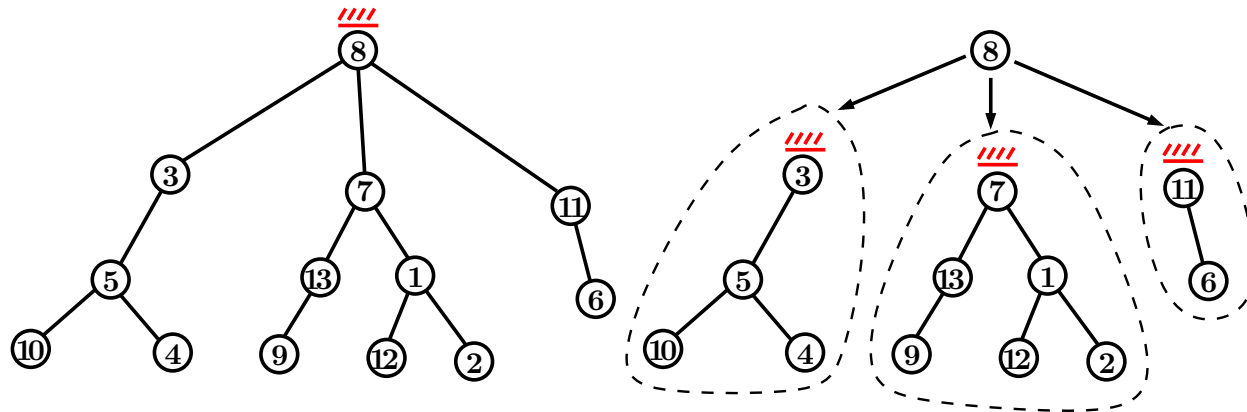
- The class  $\mathcal{R}$  of rooted trees satisfies the [decomposition](#)



$$\mathcal{R} = \mathcal{Z} \times \text{Set}(\mathcal{R}) \Rightarrow R(x) = x \exp(R(x))$$

# Rooted trees are decomposable

- The class  $\mathcal{R}$  of rooted trees satisfies the **decomposition**



$$\mathcal{R} = \mathcal{Z} \times \text{Set}(\mathcal{R}) \Rightarrow R(x) = x \exp(R(x))$$

- Lagrange **inversion formula**

inverse of  $R(x)$  is  $R^{(-1)}(y) = y \exp(-y)$

$\Rightarrow$  **Rooted trees:**  $r_n = n^{n-1} \Rightarrow$  **Trees:**  $c_n = n^{n-2}$

# Counting labelled trees: summary

- **Decomposition** of rooted trees

$$R(x) = x \exp(R(x))$$

yields  $r_n = n^{n-1}$  from Lagrange inversion formula

- **Pointing** relation:  $t_n = r_n/n$ :

$$R(x) = xT'(x)$$

yields  $t_n = n^{n-2}$

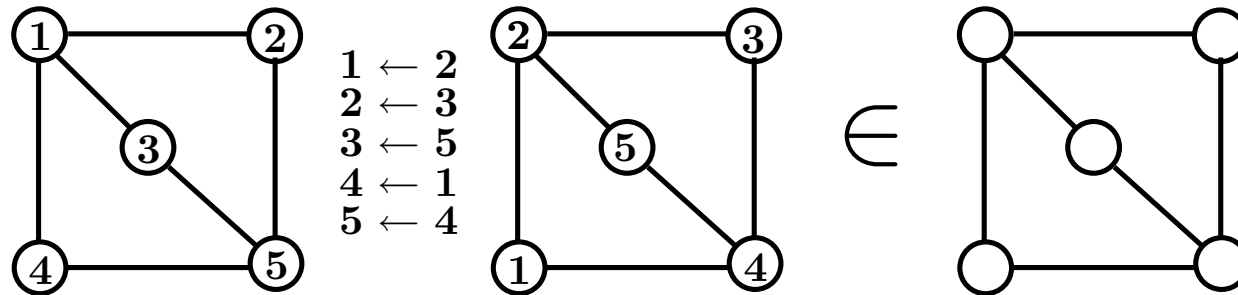
Same method applies for many classes (**planar graphs**)

# Adaptation to the unlabelled setting



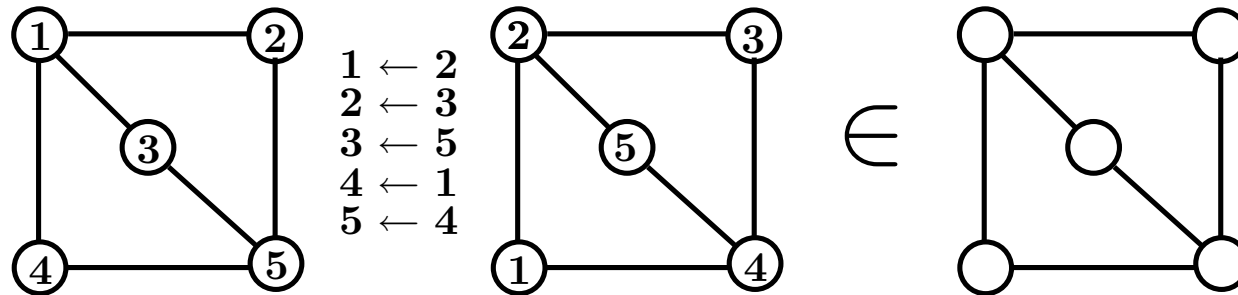
# Unlabelled setting

- Unlabelled structures = labelled structures up to relabeling

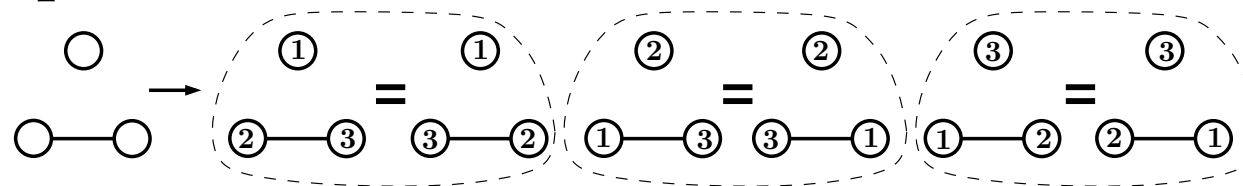


# Unlabelled setting

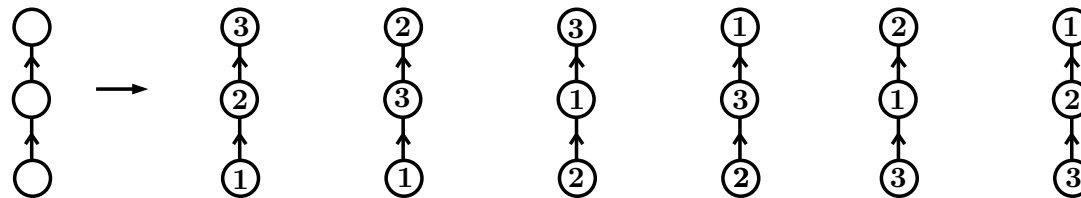
- Unlabelled structures = labelled structures up to relabeling



- Examples:



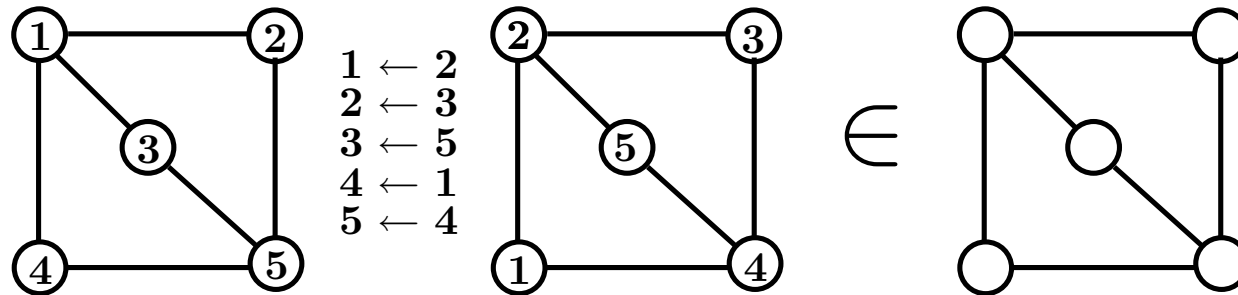
3 labeled objects (instead of  $3! = 6$ )



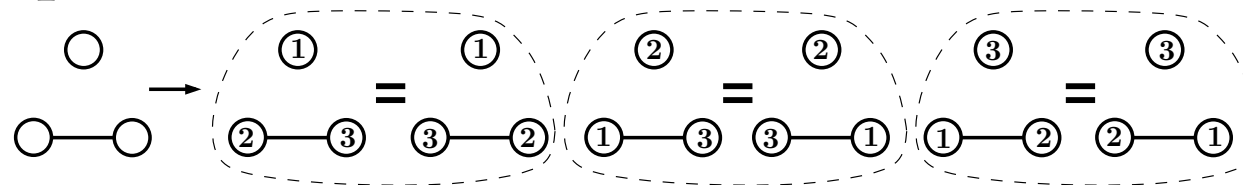
6 labeled objects (no symmetry)

# Unlabelled setting

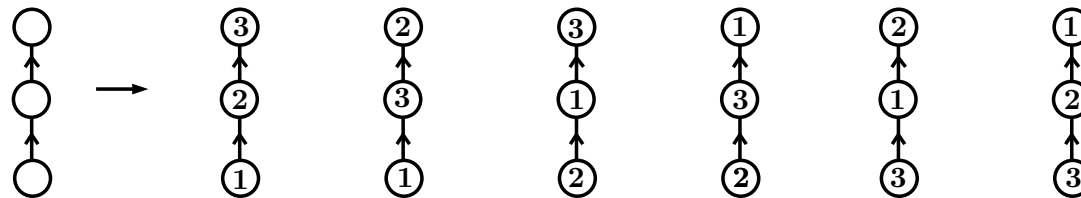
- Unlabelled structures = labelled structures up to relabeling



- Examples:



3 labeled objects (instead of  $3! = 6$ )



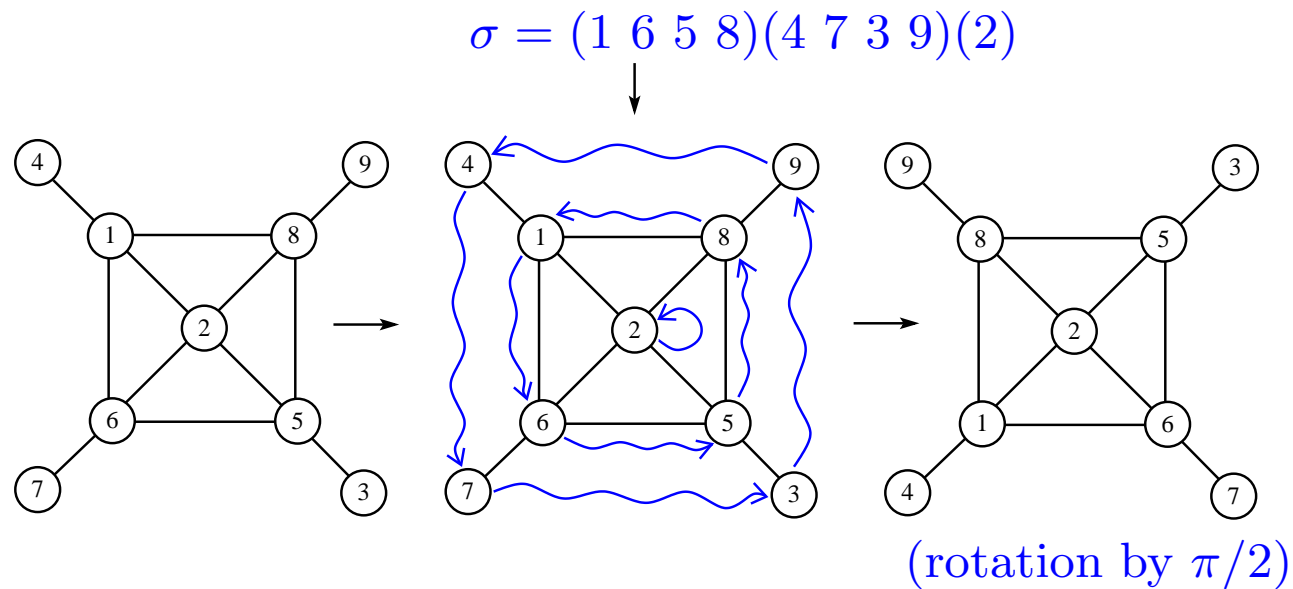
6 labeled objects (no symmetry)

- Unlabelled struct. size  $n \rightarrow$  at most  $n!$  labelled structures.  
 $\Rightarrow \frac{1}{n!} a_n^{\text{label.}} \leq a_n^{\text{unlabel.}} \Rightarrow (\text{EGF}) A(x) \preceq \tilde{A}(x) (\text{OGF}).$

# Symmetries

Let  $\mathcal{A}$  be a labelled class,

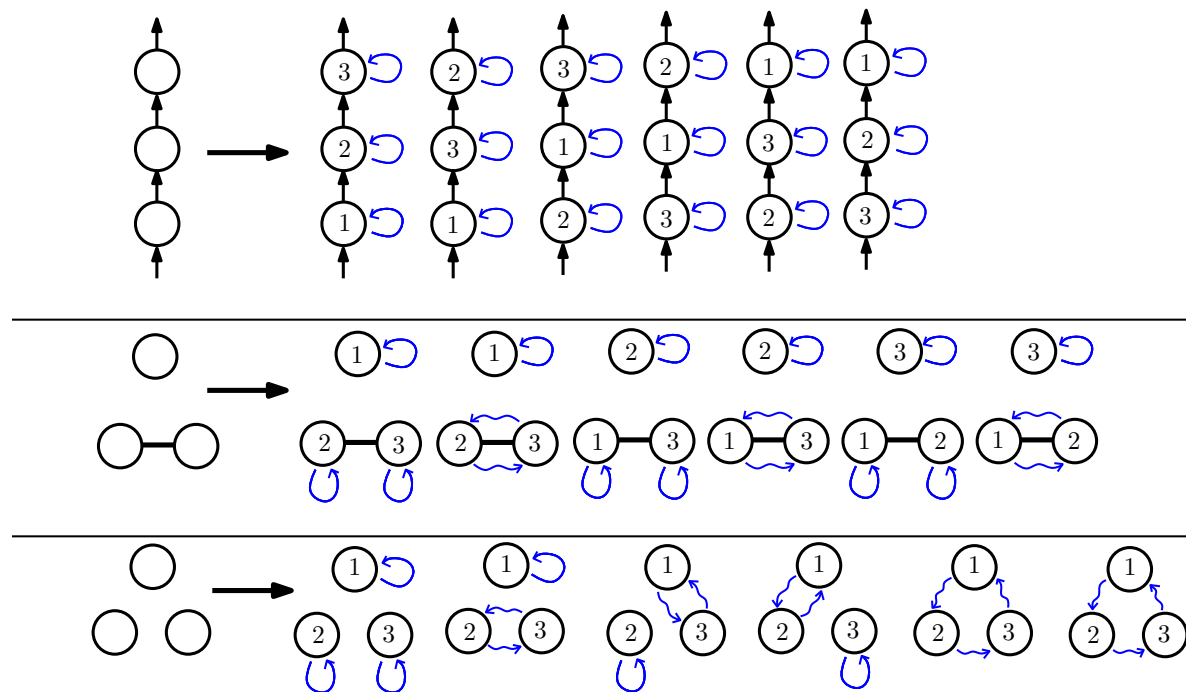
- A **symmetry** of size  $n$  on  $\mathcal{A}$  is a pair  $(\sigma \in \mathfrak{S}_n, A \in \mathcal{A}_n)$  such that  $A$  is **fixed** by the action of  $\sigma$ .



# Burnside's lemma

Given  $\mathcal{A}$  a labelled class (species of structures) let  
 $\tilde{\mathcal{A}} = \mathcal{A}/\text{isomorphisms}$ ,  $\text{Sym}(\mathcal{A}) = \{\text{Symmetries from } \mathcal{A}\}$

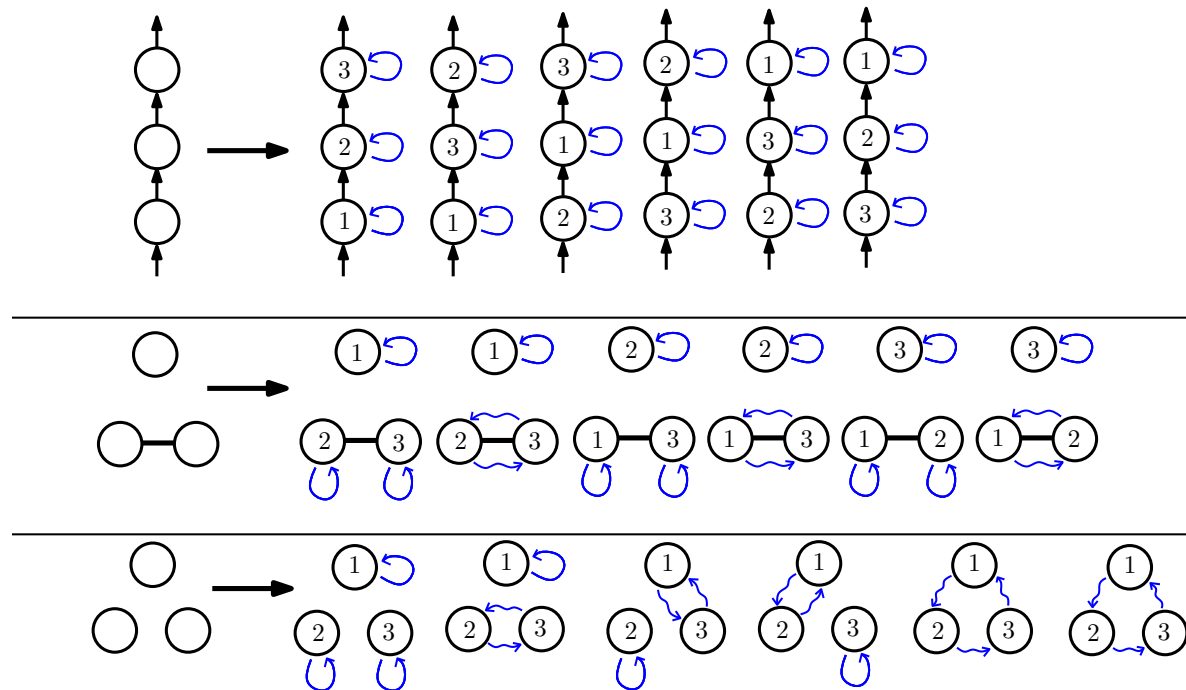
- **Burnside's lemma**  $\Rightarrow \text{Sym}(\mathcal{A})_n \simeq n! \times \tilde{\mathcal{A}}_n$



# Burnside's lemma

Given  $\mathcal{A}$  a labelled class (species of structures) let  $\tilde{\mathcal{A}} = \mathcal{A}/\text{isomorphisms}$ ,  $\text{Sym}(\mathcal{A}) = \{\text{Symmetries from } \mathcal{A}\}$

- **Burnside's lemma**  $\Rightarrow \text{Sym}(\mathcal{A})_n \simeq n! \times \tilde{\mathcal{A}}_n$



- Hence  $\text{EGF of } \text{Sym}(\mathcal{A}) = \tilde{\mathcal{A}}(x) \text{ (OGF)}$

# Cycle index sum

Let  $\mathcal{A}$  be a **labelled** class,  $\text{Sym}(\mathcal{A})$  the **symmetry** class.

- **Refined weight** for  $(\sigma, A) \in \text{Sym}(\mathcal{A})$

$$W(\sigma, A) := \frac{1}{n!} s_1^{c_1(\sigma)} s_2^{c_2(\sigma)} \dots s_n^{c_n(\sigma)}$$

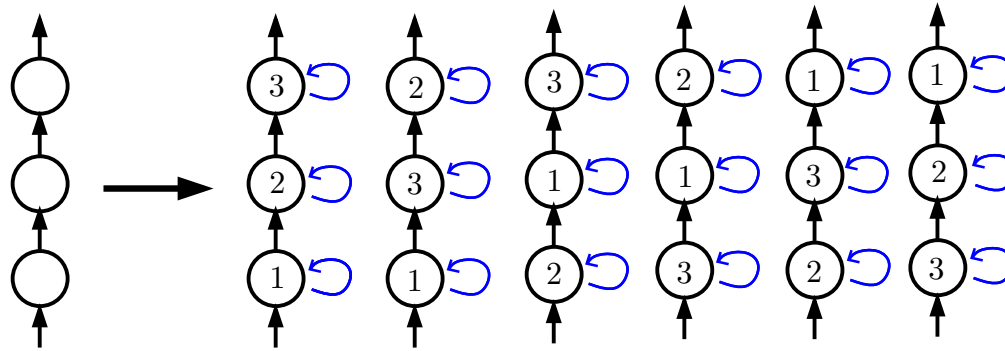
where  $c_i(\sigma) = \#\{\text{cycles length } i \text{ in } \sigma\}$

- Cycle index sum of  $\mathcal{A}$  (cf Pólya) is the **multivariate** series

$$\begin{aligned} Z_{\mathcal{A}}(s_1, s_2, \dots) &= \sum_{\sigma \cdot A = A} W(\sigma, A) \\ &= \sum_{n \geq 1} \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} s_1^{c_1} \dots s_n^{c_n} \#(\text{Fix}_{\sigma}) \end{aligned}$$

- OGF of  $\tilde{\mathcal{A}} = \text{EGF of } \text{Sym}(\mathcal{A}) = Z_{\mathcal{A}}(x, x^2, x^3, \dots)$

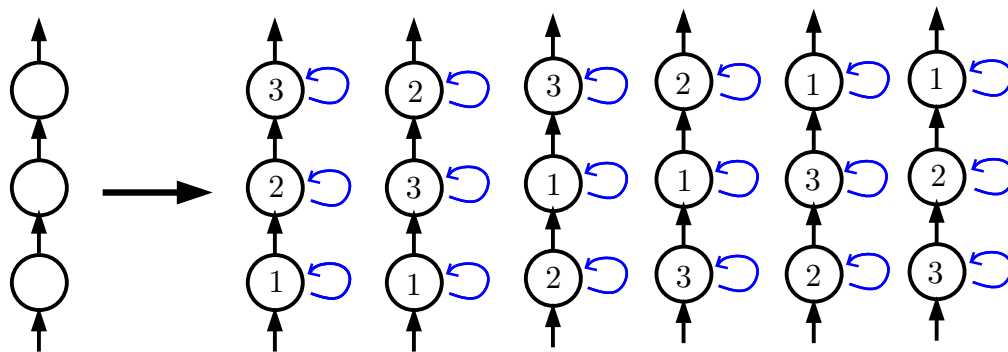
# Examples of cycle index sums



$$\begin{aligned}
 Z &= \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 \\
 &= s_1^3
 \end{aligned}$$

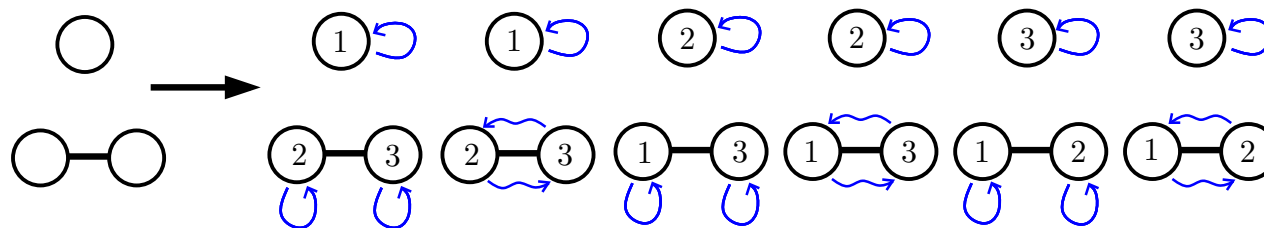


# Examples of cycle index sums



$$Z = \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3$$

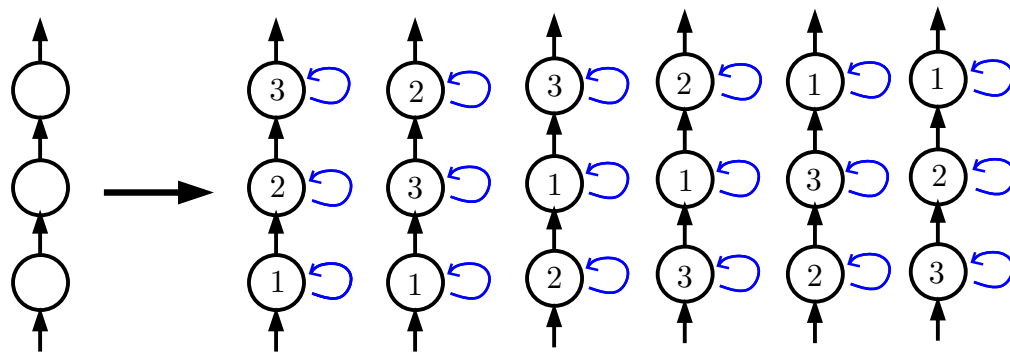
$$= s_1^3$$



$$Z = \frac{1}{6} s_1^3 + \frac{1}{6} s_1 s_2 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1 s_2 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1 s_2$$

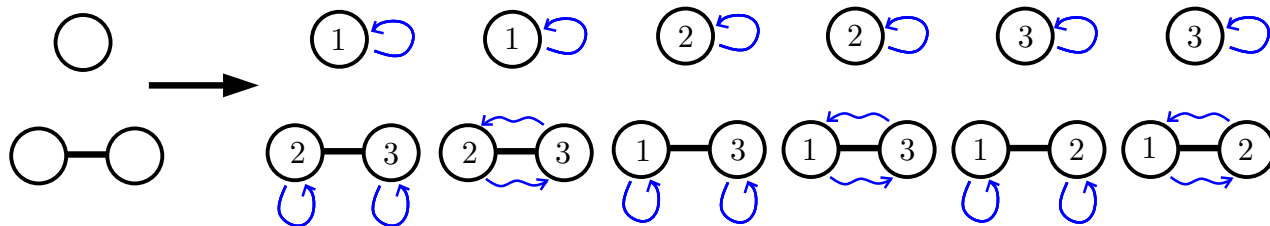
$$= \frac{1}{2} s_1^3 + \frac{1}{2} s_1 s_2$$

# Examples of cycle index sums



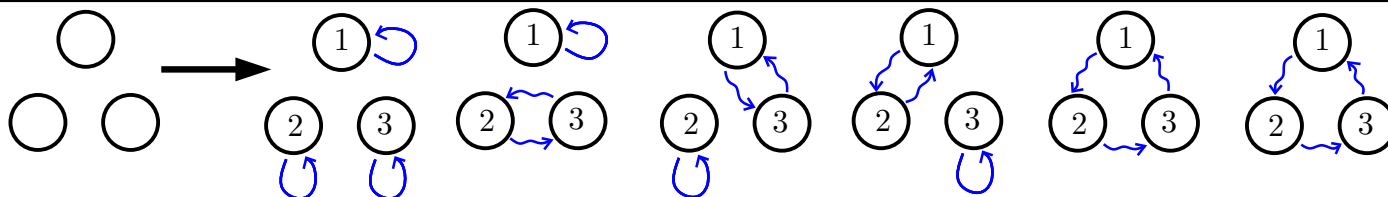
$$Z = \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1^3$$

$$= s_1^3$$



$$Z = \frac{1}{6} s_1^3 + \frac{1}{6} s_1 s_2 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1 s_2 + \frac{1}{6} s_1^3 + \frac{1}{6} s_1 s_2$$

$$= \frac{1}{2} s_1^3 + \frac{1}{2} s_1 s_2$$



$$Z = \frac{1}{6} s_1^3 + \frac{1}{6} s_1 s_2 + \frac{1}{6} s_1 s_2 + \frac{1}{6} s_1 s_2 + \frac{1}{6} s_3 + \frac{1}{6} s_3$$

# Dictionary for OGF

-

# Dictionary for OGF

- **Unlabelled** class  $\tilde{\mathcal{C}} = \cup_n \mathcal{C}_n / \mathfrak{S}_n$     $\tilde{c}_n = \text{Card}(\tilde{\mathcal{C}}_n)$

$$\text{OGF : } \tilde{C}(x) = \sum_{n \geq 0} \tilde{c}_n x^n$$

- **Dictionary** (computation rules):

Disjoint union	$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$\tilde{C}(x) = \tilde{A}(x) + \tilde{B}(x)$
Product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$\tilde{C}(x) = \tilde{A}(x)\tilde{B}(x)$
Set	$\mathcal{C} = \text{Set}(\mathcal{A})$	$\tilde{C}(x) = \exp\left(\sum_{k \geq 1} \frac{1}{k} \tilde{A}(x^k)\right)$
Substitution	$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$	$\tilde{C}(x) = Z_{\mathcal{A}}(\tilde{B}(x), \tilde{B}(x^2), \dots)$

# Dictionary for OGF

- **Unlabelled** class  $\tilde{\mathcal{C}} = \cup_n \mathcal{C}_n / \mathfrak{S}_n$     $\tilde{c}_n = \text{Card}(\tilde{\mathcal{C}}_n)$

$$\text{OGF : } \tilde{C}(x) = \sum_{n \geq 0} \tilde{c}_n x^n$$

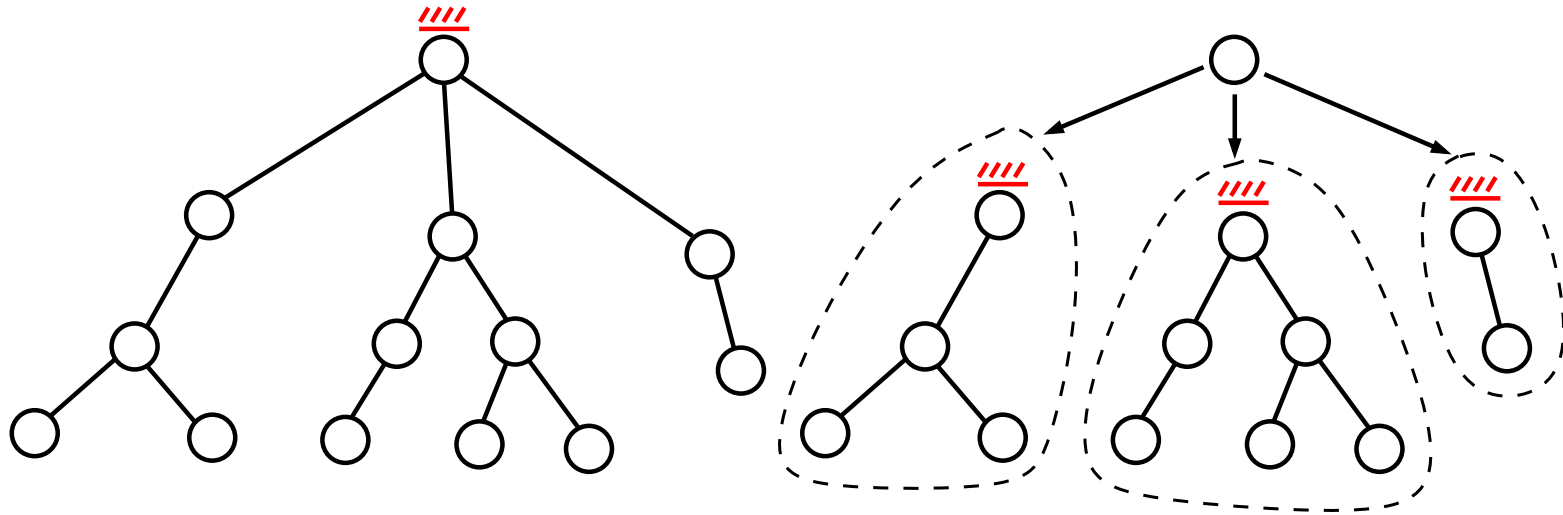
- **Dictionary** (computation rules):

Disjoint union	$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$\tilde{C}(x) = \tilde{A}(x) + \tilde{B}(x)$
Product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$\tilde{C}(x) = \tilde{A}(x)\tilde{B}(x)$
Set	$\mathcal{C} = \text{Set}(\mathcal{A})$	$\tilde{C}(x) = \exp\left(\sum_{k \geq 1} \frac{1}{k} \tilde{A}(x^k)\right)$
Substitution	$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$	$\tilde{C}(x) = Z_{\mathcal{A}}(\tilde{B}(x), \tilde{B}(x^2), \dots)$

- **Remark:**  $\text{Set}(\mathcal{A}) = \text{Set} \circ \mathcal{A}$ , computation rule for  $\circ$  implies the one for Set using  $Z_{\text{Set}} = \exp(\sum_{i \geq 1} s_i / i)$

# Example: rooted trees

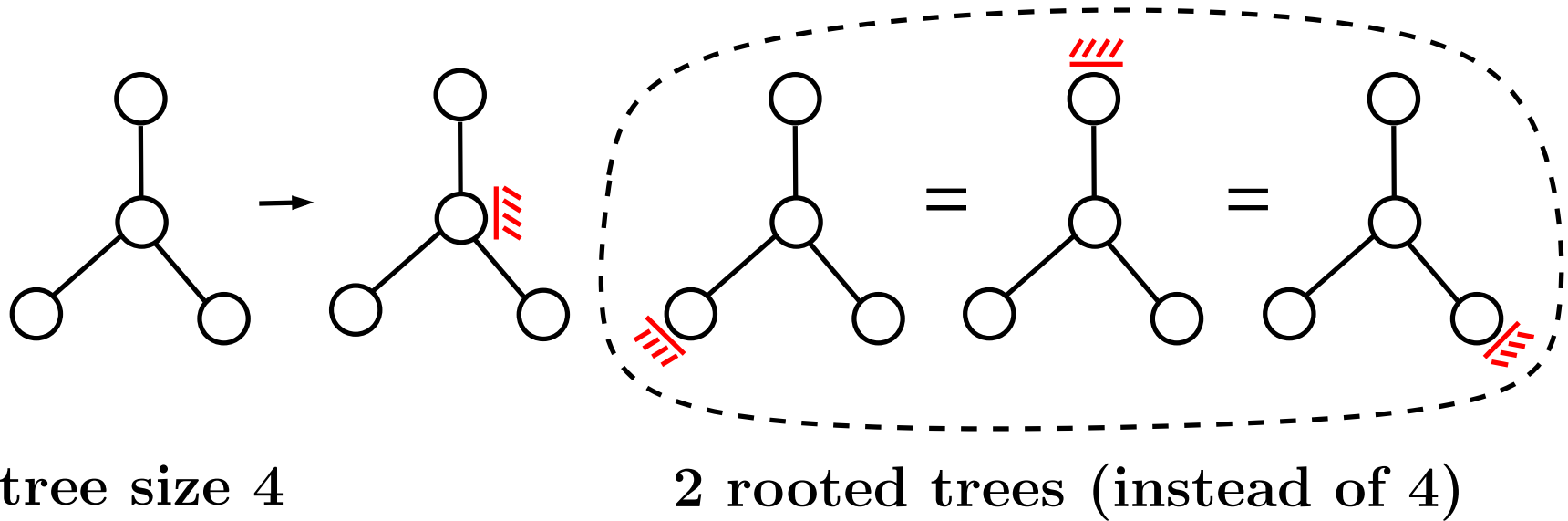
Decomposition at the root:



$$\mathcal{R} = \mathcal{Z} \times \text{Set}(\mathcal{R}) \Rightarrow R(z) = z \exp \left( \sum_{i \geq 1} \frac{1}{i} R(z^i) \right)$$

$\Rightarrow$  **recurrence** formula for  $[z^n]R(z)$

# Count rooted $\not\Rightarrow$ count unrooted

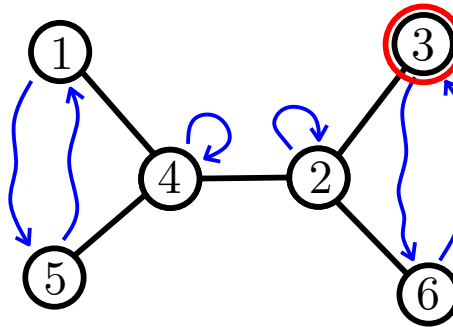


In general  $n \cdot a_n^{\text{unrooted}} > a_n^{\text{rooted}}$  (symmetries)

**Question:**  $n \cdot a_n^{\text{unrooted}} = ?$

# Pointing symmetries

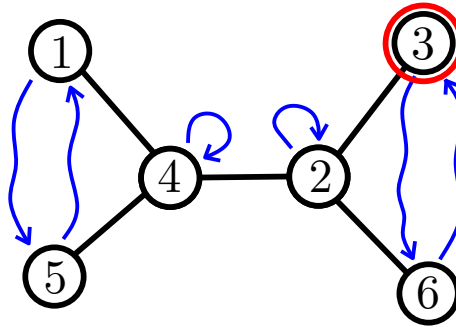
Pointed symmetry = symmetry + marked atom





# Pointing symmetries

Pointed symmetry = symmetry + marked atom

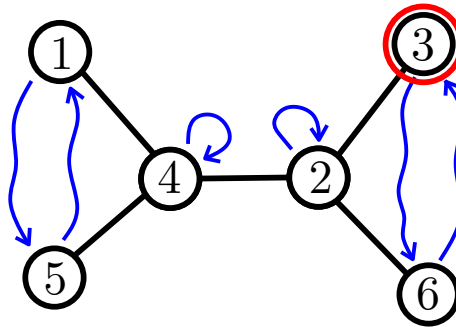


In size  $n$  we have

$$\begin{array}{ccc} \tilde{\mathcal{A}} & \xrightarrow{\times n!} & \text{Sym}(\mathcal{A}) \\ & & \downarrow \times n \\ & & (\text{Sym}(\mathcal{A}))^\bullet \end{array}$$

# Pointing symmetries

Pointed symmetry = symmetry + marked atom



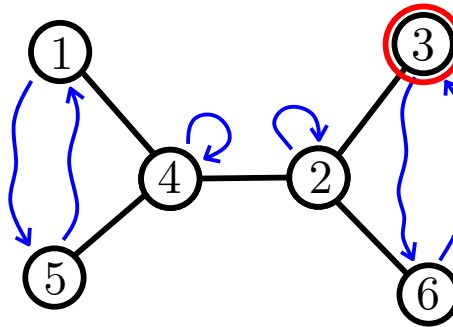
In size  $n$  we have

$$\begin{array}{ccc}
 \tilde{\mathcal{A}} & \xrightarrow{\times n!} & \text{Sym}(\mathcal{A}) \\
 & & \downarrow \times n \\
 & & (\text{Sym}(\mathcal{A}))^\bullet
 \end{array}$$

Look for a class  $\mathcal{P}$  such that  $\text{Sym}(\mathcal{P}) \simeq (\text{Sym}(\mathcal{A}))^\bullet$

# Pointing symmetries

Pointed symmetry = symmetry + marked atom



In size  $n$  we have

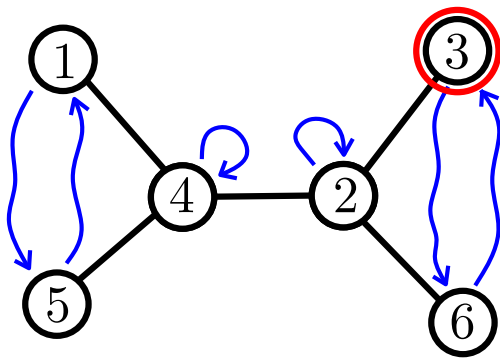
$$\begin{array}{ccc}
 \tilde{\mathcal{A}} & \xrightarrow{\times n!} & \text{Sym}(\mathcal{A}) \\
 & & \downarrow \times n \\
 & & (\text{Sym}(\mathcal{A}))^\bullet
 \end{array}$$

Look for a class  $\mathcal{P}$  such that  $\text{Sym}(\mathcal{P}) \simeq (\text{Sym}(\mathcal{A}))^\bullet$

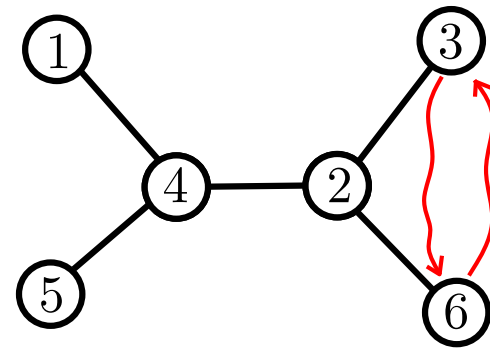
(rk:  $\text{Sym}(\mathcal{A}^\bullet) \subseteq (\text{Sym}(\mathcal{A}))^\bullet$ )

# Cycle-pointed structures

**Definition:** Cycle-pointed structure = structure  $A$  + cycle  $c$  such that there exists (at least) one automorphism of  $A$  having  $c$  as one of its cycles.



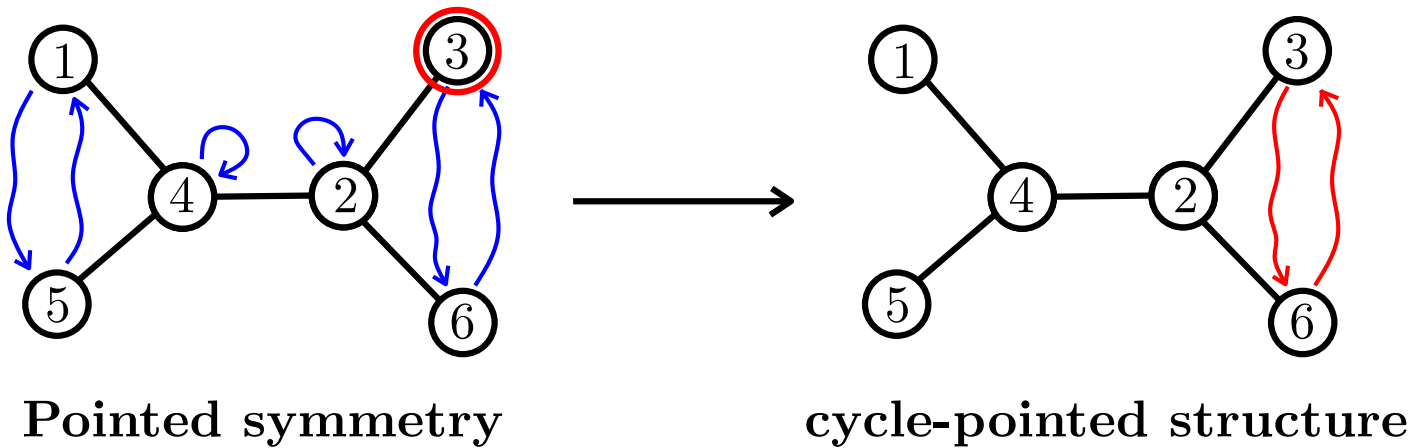
Pointed symmetry



cycle-pointed structure

# Cycle-pointed structures

**Definition:** Cycle-pointed structure = structure  $A$  + cycle  $c$  such that there exists (at least) one automorphism of  $A$  having  $c$  as one of its cycles.



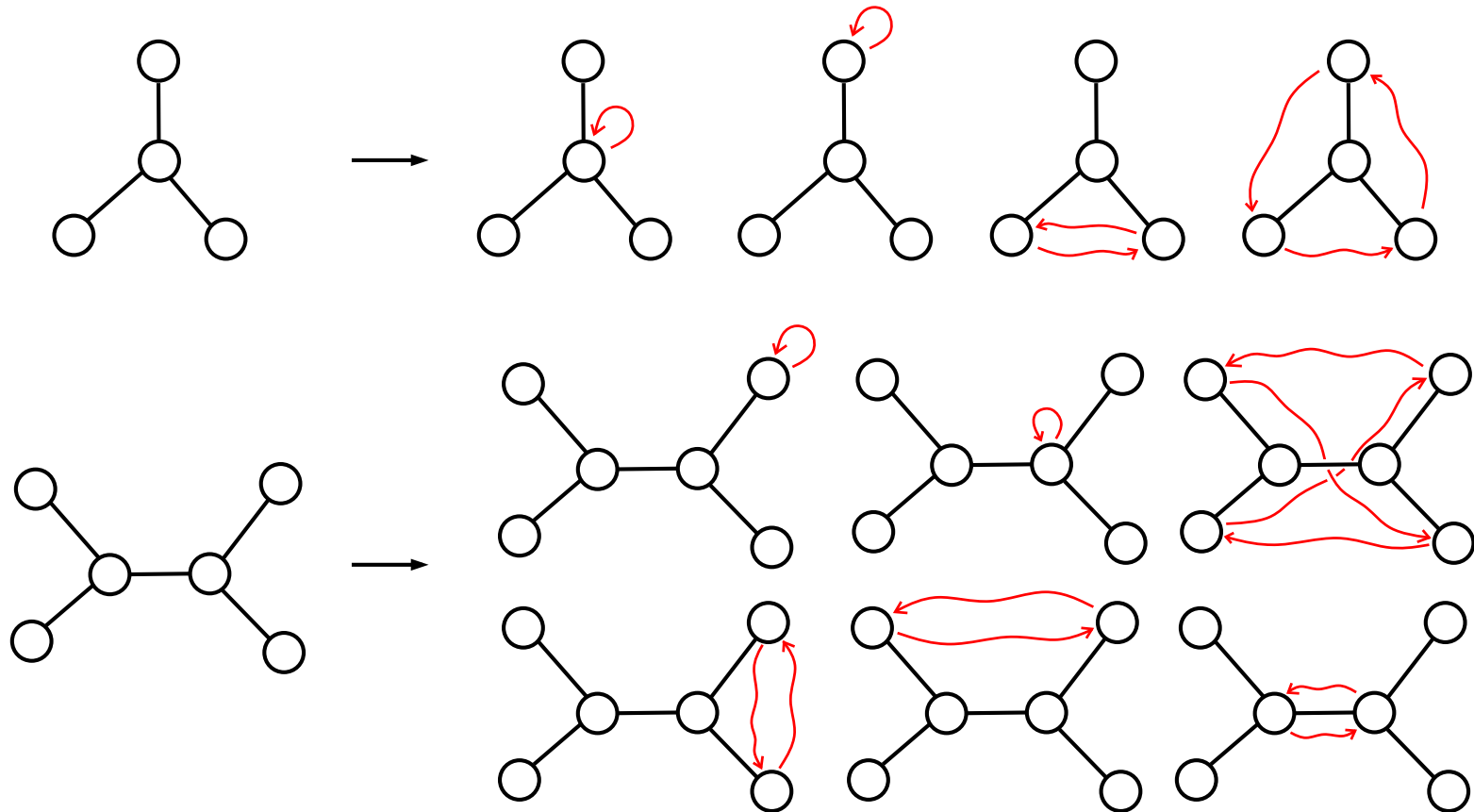
Let  $\mathcal{A}^\circ = \{\text{cycle - pointed structures from } \mathcal{A}\}$ . Then

$$\text{Sym}(\mathcal{A}^\circ) \simeq (\text{Sym}(\mathcal{A}))^\bullet$$

# Cycle-pointing is unbiased

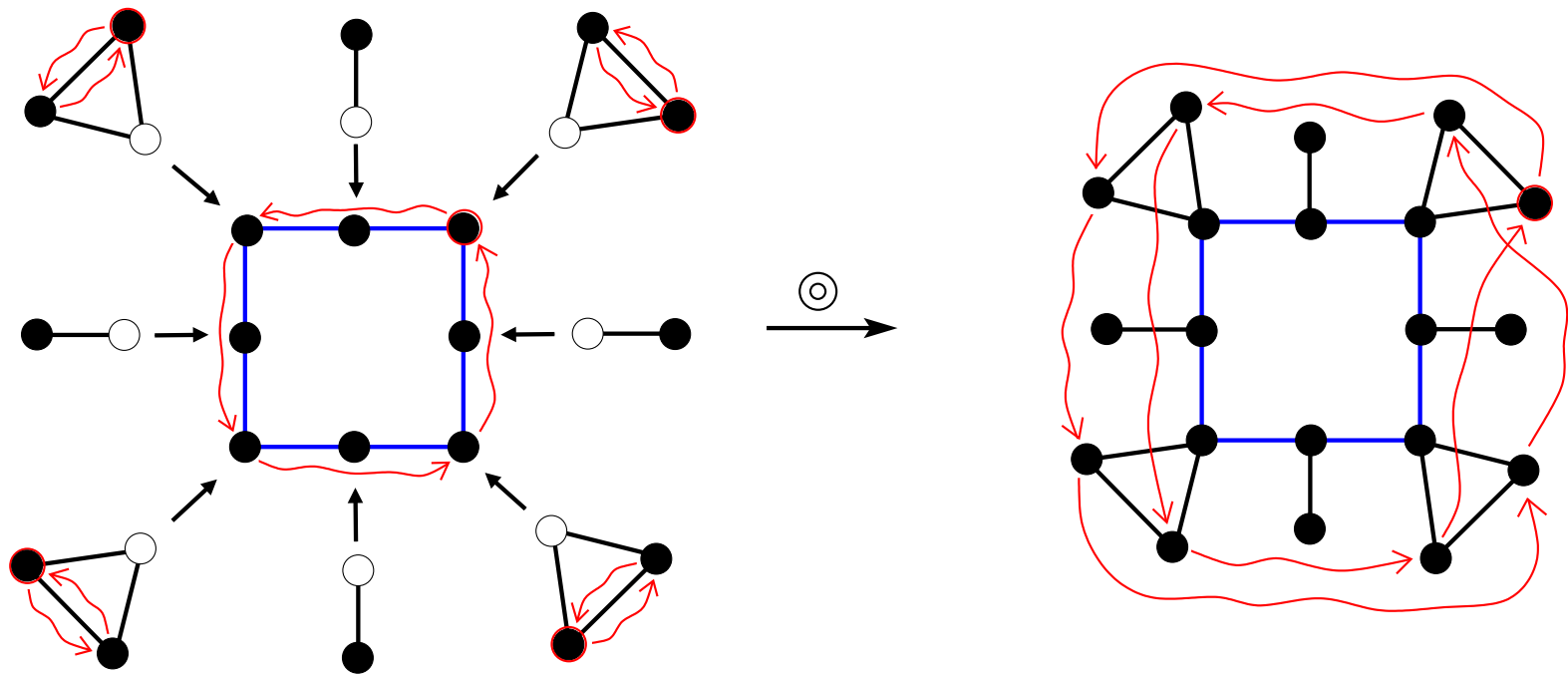
**Theorem:** An unlabelled structure of size  $n$  gives rise to  $n$  unlabelled cycle-pointed structures (cf Parker's lemma).

$$\widetilde{\mathcal{A}}_n^\circ \cong n \times \widetilde{\mathcal{A}}_n$$



# Pointing the classical constructions

- $(\mathcal{A} + \mathcal{B})^\circ = \mathcal{A}^\circ + \mathcal{B}^\circ$
- $(\mathcal{A} \times \mathcal{B})^\circ = \mathcal{A}^\circ \times \mathcal{B} + \mathcal{A} \times \mathcal{B}^\circ$
- $(\mathcal{A} \circ \mathcal{B})^\circ = \mathcal{A}^\circ \odot \mathcal{B}$



# Application: counting trees (1)

## Decomposition of cycle-pointed trees (3 lines)

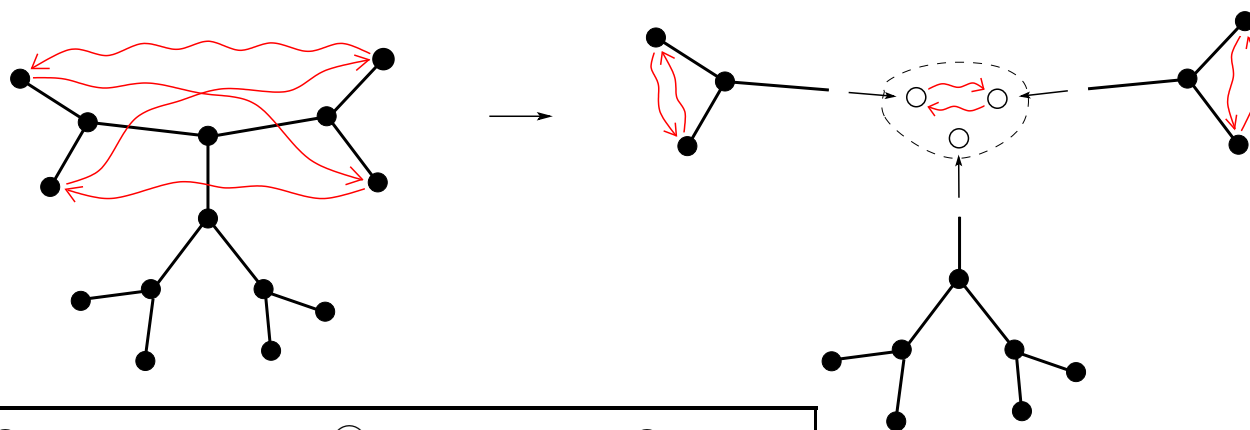
- Dichotomy: pointed cycle has length 1 or  $\geq 2$ :

$$1) \mathcal{T}^\circ = \mathcal{R} + \mathcal{T}^*$$

- Rooted trees ( $\mathcal{R}$ ) are decomposed at the root

$$2) \mathcal{R} = \mathcal{Z} \times \text{Set}(\mathcal{R})$$

- Symmetric cycle-pointed trees ( $\mathcal{T}^*$ ) are decomposed at a centre of symmetry.



$$3) \mathcal{T}^* = \mathcal{Z} \times \text{Set}^* \odot \mathcal{R} + \mathcal{L}^* \odot \mathcal{R}$$



# Application: counting trees (2)

$$\begin{cases} T^\circ &= \mathcal{R} + T^{\circ*} \\ \mathcal{R} &= \mathcal{Z} \times \text{Set}(\mathcal{R}) \\ T^{\circ*} &= \mathcal{Z} \times \text{Set}^{\circ*} \odot (\mathcal{R}) + \mathcal{L}^{\circ*} \odot (\mathcal{R}) \end{cases}$$

↓ translate to equation system  
(dictionary rules+Pólya operators)

$$\begin{cases} R(x) &= x \exp \left( \sum_{k \geq 1} \frac{1}{k} R(x^k) \right) \\ xt'(x) &= R(x) + x^2 R'(x^2) + \sum_{i \geq 2} x^i R'(x^i) R(x) \end{cases}$$

↓ extract coefficients

$n$	1	2	3	4	5	6	7	8	9
$\widetilde{a_n^\circ}$	1	2	3	8	15	36	77	184	423
$\widetilde{a_n}$	1	1	1	2	3	6	11	23	47

# Application: counting trees (2)

$$\begin{cases} \mathcal{T}^\circ &= \mathcal{R} + \mathcal{T}^* \\ \mathcal{R} &= \mathcal{Z} \times \text{Set}(\mathcal{R}) \\ \mathcal{T}^* &= \mathcal{Z} \times \text{Set}^* \odot (\mathcal{R}) + \mathcal{L}^* \odot (\mathcal{R}) \end{cases}$$

↓ translate to equation system  
(dictionary rules+Pólya operators)

$$\begin{cases} R(x) &= x \exp \left( \sum_{k \geq 1} \frac{1}{k} R(x^k) \right) \\ xt'(x) &= R(x) + x^2 R'(x^2) + \sum_{i \geq 2} x^i R'(x^i) R(x) \\ &= xR'(x)(1 - R(x)) + x^2 R'(x^2) \end{cases}$$

↓ extract coefficients

$n$	1	2	3	4	5	6	7	8	9
$\widetilde{a_n^\circ}$	1	2	3	8	15	36	77	184	423
$\widetilde{a_n}$	1	1	1	2	3	6	11	23	47

# Exact counting results

**Theorem:** [Bergeron, Labelle, Leroux],

For any class  $\mathcal{A}$  **decomposable** in terms of

- basic classes  $\{\mathbf{1}, \mathcal{Z}, \text{Seq}, \text{Cyc}, \text{Set}\}$ ,
- constructions  $\{+, \times, \circ\}$

the counting coefficients  $|\widetilde{\mathcal{A}}_n|$  **can be computed automatically**

# Exact counting results

**Theorem:** [Bergeron, Labelle, Leroux], [Bodirsky et al'07,10]

For any class  $\mathcal{A}$  **decomposable** in terms of

- basic classes  $\{\mathbf{1}, \mathcal{Z}, \text{Seq}, \text{Cyc}, \text{Set}\}$ ,
- constructions  $\{+, \times, \circ\}$  **and**  $\odot$ ,
- **cycle-pointing operator**  $\mathcal{C} \mapsto \mathcal{C}^\circ$ ,

the counting coefficients  $|\widetilde{\mathcal{A}}_n|$  **can be computed automatically**

# Exact counting results

**Theorem:** [Bergeron, Labelle, Leroux], [Bodirsky et al'07,10]

For any class  $\mathcal{A}$  **decomposable** in terms of

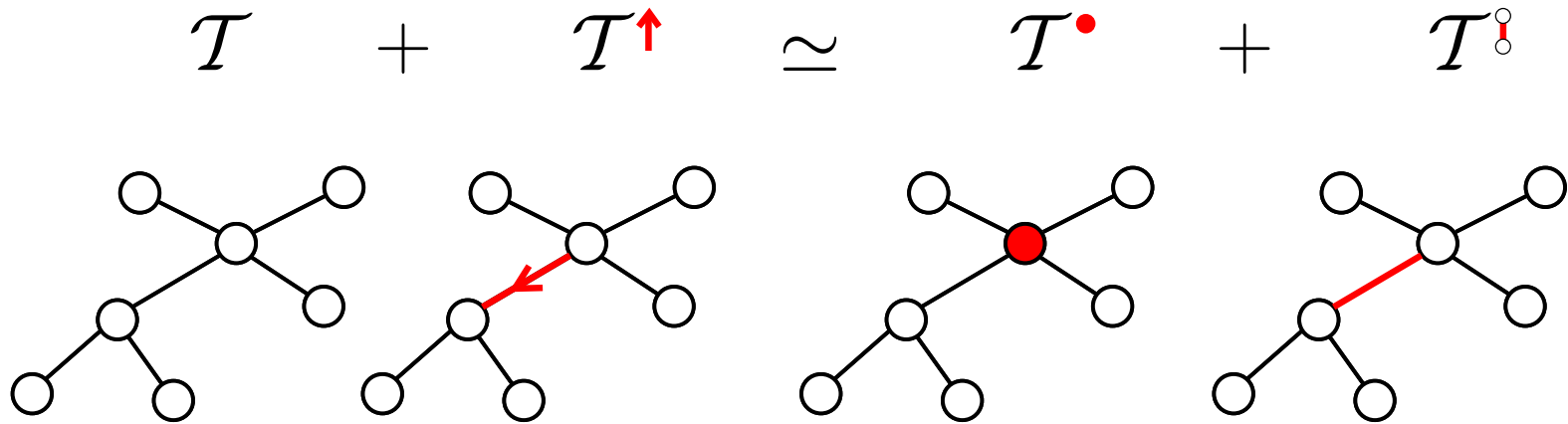
- basic classes  $\{\mathbf{1}, \mathcal{Z}, \text{Seq}, \text{Cyc}, \text{Set}\}$ ,
- constructions  $\{+, \times, \circ\}$  **and**  $\odot$ ,
- **cycle-pointing operator**  $\mathcal{C} \mapsto \mathcal{C}^\circ$ ,

the counting coefficients  $|\widetilde{\mathcal{A}}_n|$  **can be computed automatically**

(includes tree families, outerplanar graphs,...)

# Another approach

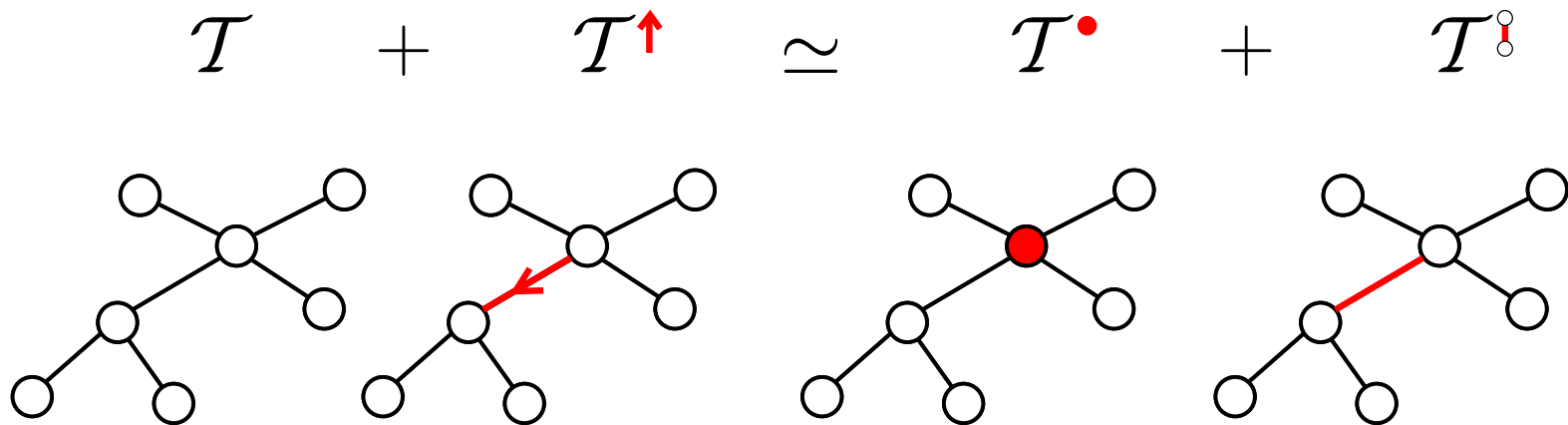
- Dissimilarity characteristic formula (Otter)
- Dissymmetry theorem (Robinson, Leroux):



$$\mathbf{t}(\mathbf{x}) = \mathbf{t}^{\bullet}(\mathbf{x}) + \mathbf{t}^{\circ}(\mathbf{x}) - \mathbf{t}^{\uparrow}(\mathbf{x})$$

# Another approach

- Dissimilarity characteristic formula (Otter)
- Dissymmetry theorem (Robinson, Leroux):

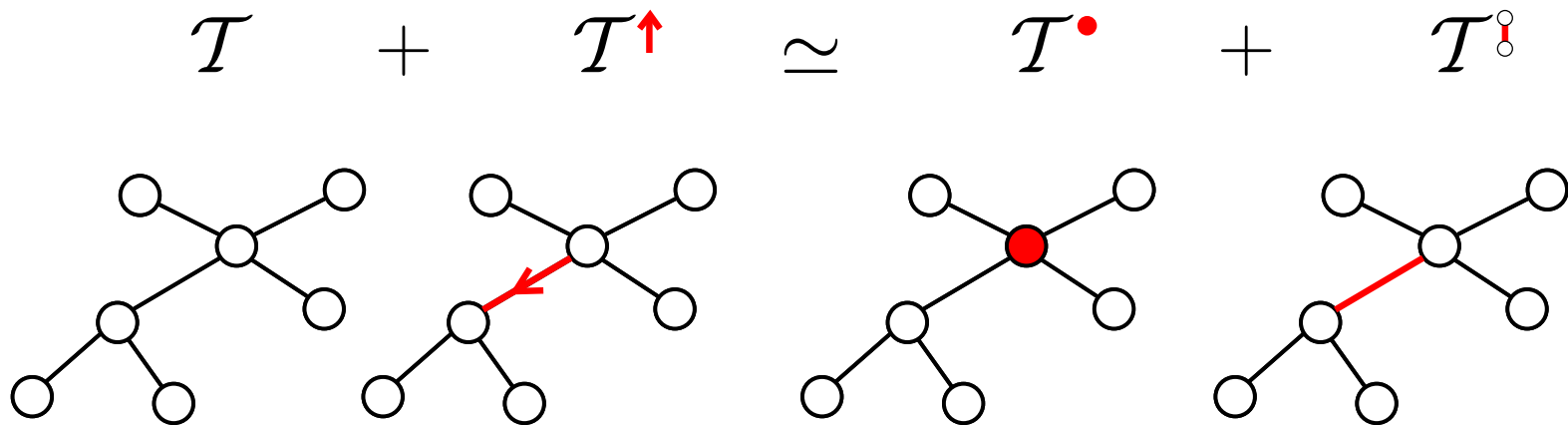


$$\mathbf{t}(\mathbf{x}) = \mathbf{t}^{\bullet}(\mathbf{x}) + \mathbf{t}^{\circ}(\mathbf{x}) - \mathbf{t}^{\uparrow}(\mathbf{x})$$

$$\Rightarrow t(x) = R(x) - (R(x)^2 - R(x^2)) \text{ with } R(x) = \exp\left(\sum_{i \geq 1} \frac{1}{i} R(x^i)\right)$$

# Another approach

- Dissimilarity characteristic formula (Otter)
- Dissymmetry theorem (Robinson, Leroux):



$$\mathbf{t}(\mathbf{x}) = \mathbf{t}^{\bullet}(\mathbf{x}) + \mathbf{t}^{\circ}(\mathbf{x}) - \mathbf{t}^{\uparrow}(\mathbf{x})$$

$$\Rightarrow t(x) = R(x) - (R(x)^2 - R(x^2)) \text{ with } R(x) = \exp\left(\sum_{i \geq 1} \frac{1}{i} R(x^i)\right)$$

$$(\text{agrees with } xt'(x) = xR'(x)(1 - R(x)) + x^2R'(x^2))$$



# Application to asymptotic enumeration

# Asymptotic scheme

Main result: “universality” of asymptotic behaviour

$$|\widetilde{\mathcal{A}}_n| \sim c \gamma^n n^{-5/2}$$

for “any” unrooted “tree-like” family  $\mathcal{A}$

# Asymptotic scheme

Main result: “universality” of asymptotic behaviour

$$|\widetilde{\mathcal{A}}_n| \sim c \gamma^n n^{-5/2}$$

for “any” unrooted “tree-like” family  $\mathcal{A}$

Scheme:

- Decompose cycle-pointed class  $\mathcal{A}^\circ$   
 $\Rightarrow$  Equation for  $\widetilde{A}^\circ(x)$
- Drmota-Lalley-Woods  $\Rightarrow \widetilde{A}^\circ(x)$  has square-root sing.
- Transfer theorem [Flajolet-Odlyzko]  $\Rightarrow |\widetilde{\mathcal{A}}_n^\circ| \sim c \gamma^n n^{-3/2}$
- Pointing relation:  $|\widetilde{\mathcal{A}}_n| = \frac{1}{n} |\widetilde{\mathcal{A}}_n^\circ| \sim c \gamma^n n^{-5/2}$

# Illustration on trees

- Rooted labelled trees:  $y = L(z)$  satisfies

$$y = z \exp(y)$$

Inverse is  $g(y) = y \exp(-y)$ ,

$$g'(y) = 0 \Rightarrow y = 1 \Rightarrow z = 1/e \Rightarrow L(z) = 1 - c\sqrt{1 - ze} + \dots$$

# Illustration on trees

- **Rooted labelled trees:**  $y = L(z)$  satisfies

$$y = z \exp(y)$$

Inverse is  $g(y) = y \exp(-y)$ ,

$$g'(y) = 0 \Rightarrow y = 1 \Rightarrow z = 1/e \Rightarrow L(z) = 1 - c\sqrt{1 - ze} + \dots$$

- **Rooted unlabelled trees:**  $y = R(z)$  satisfies

$$y = z \exp(y + A(z)), \quad \text{where } A(z) = \sum_{i \geq 2} \frac{1}{i} R(z^i)$$

# Illustration on trees

- **Rooted labelled trees:**  $y = L(z)$  satisfies

$$y = z \exp(y)$$

Inverse is  $g(y) = y \exp(-y)$ ,

$$g'(y) = 0 \Rightarrow y = 1 \Rightarrow z = 1/e \Rightarrow L(z) = 1 - c\sqrt{1 - ze} + \dots$$

- **Rooted unlabelled trees:**  $y = R(z)$  satisfies

$$y = z \exp(y + A(z)), \quad \text{where } A(z) = \sum_{i \geq 2} \frac{1}{i} R(z^i)$$

Hence  $R(z) = L(A(z)) = 1 - c'\sqrt{1 - z/\rho} + \dots$

where  $\rho$  satisfies  $A(\rho) = 1/e$

# Illustration on trees

- **Rooted labelled trees:**  $y = L(z)$  satisfies

$$y = z \exp(y)$$

Inverse is  $g(y) = y \exp(-y)$ ,

$$g'(y) = 0 \Rightarrow y = 1 \Rightarrow z = 1/e \Rightarrow L(z) = 1 - c\sqrt{1 - ze} + \dots$$

- **Rooted unlabelled trees:**  $y = R(z)$  satisfies

$$y = z \exp(y + A(z)), \quad \text{where } A(z) = \sum_{i \geq 2} \frac{1}{i} R(z^i)$$

Hence  $R(z) = L(A(z)) = 1 - c'\sqrt{1 - z/\rho} + \dots$

where  $\rho$  satisfies  $A(\rho) = 1/e$

- **Cycle-pointed trees**

$$zt'(z) = \underbrace{z^2 R'(z^2)}_{\text{analytic at } \rho} + \underbrace{\left(1 + \sum_{i \geq 2} z^i R'(z^i)\right)}_{G(z) \text{ analytic at } \rho} \cdot R(z)$$

# Illustration on trees

- **Rooted labelled trees:**  $y = L(z)$  satisfies

$$y = z \exp(y)$$

Inverse is  $g(y) = y \exp(-y)$ ,

$$g'(y) = 0 \Rightarrow y = 1 \Rightarrow z = 1/e \Rightarrow L(z) = 1 - c\sqrt{1 - ze} + \dots$$

- **Rooted unlabelled trees:**  $y = R(z)$  satisfies

$$y = z \exp(y + A(z)), \quad \text{where } A(z) = \sum_{i \geq 2} \frac{1}{i} R(z^i)$$

Hence  $R(z) = L(A(z)) = 1 - c'\sqrt{1 - z/\rho} + \dots$

where  $\rho$  satisfies  $A(\rho) = 1/e$

- **Cycle-pointed trees**

$$zt'(z) = \underbrace{z^2 R'(z^2)}_{\text{analytic at } \rho} + \underbrace{\left(1 + \sum_{i \geq 2} z^i R'(z^i)\right)}_{G(z) \text{ analytic at } \rho} \cdot R(z)$$

(Rk:  $\mathbb{E}_n(\#\text{dissimilar vertices in tree}) \sim n/G(\rho)$ )



# Asymptotic using dissym. theorem

$$t(z) = R(z) - \frac{1}{2}(R(z)^2 - R(z^2))$$

Square-root expansion:  $R(z) = 1 - * \sqrt{1 - z/\rho} + \dots$

$\Rightarrow$  square-root terms cancel out,  $t(z)$  “ $\leq$ ”  $(1 - z/\rho)^{3/2}$

But  $zt'(z) \geq R(z)$ , so  $t(z)$  “ $\geq$ ”  $(1 - z/\rho)^{3/2}$

Hence (transfer theorem):

$$[z^n]t(z) \sim c \rho^{-n} n^{-5/2}$$

(Rk: Cancellation proof uneasy if “big” functional equation)

# Application to random generation

# General methods

Two sampling methods giving uniform distribution

- **Recursive method** (Nijenhuis-Wilf'78):

$$\mathbb{P}(\gamma \in \mathcal{C}_n) = \frac{1}{c_n} \quad [\text{Fixed size}]$$

- Based on **coefficients**:

$$\mathcal{C} = \mathcal{A} + \mathcal{B} \Rightarrow \mathbb{P}(\gamma \in \mathcal{A}_n) = \frac{a_n}{c_n}$$

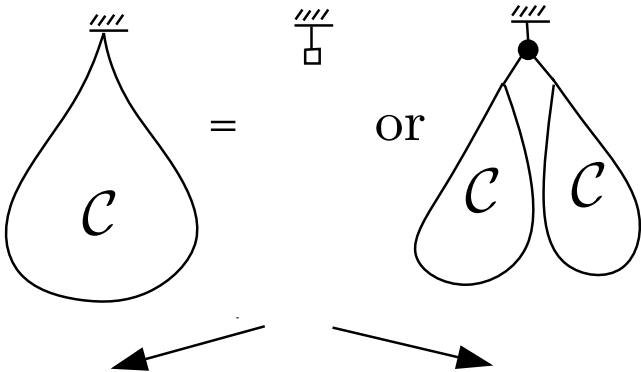
- **Boltzmann samplers** (Duchon, Flajolet, Louchard, Schaeffer'02)

$$\mathbb{P}(\gamma \in \mathcal{C}) = \frac{x^{|\gamma|}}{C(x)} \quad [\text{Whole class}]$$

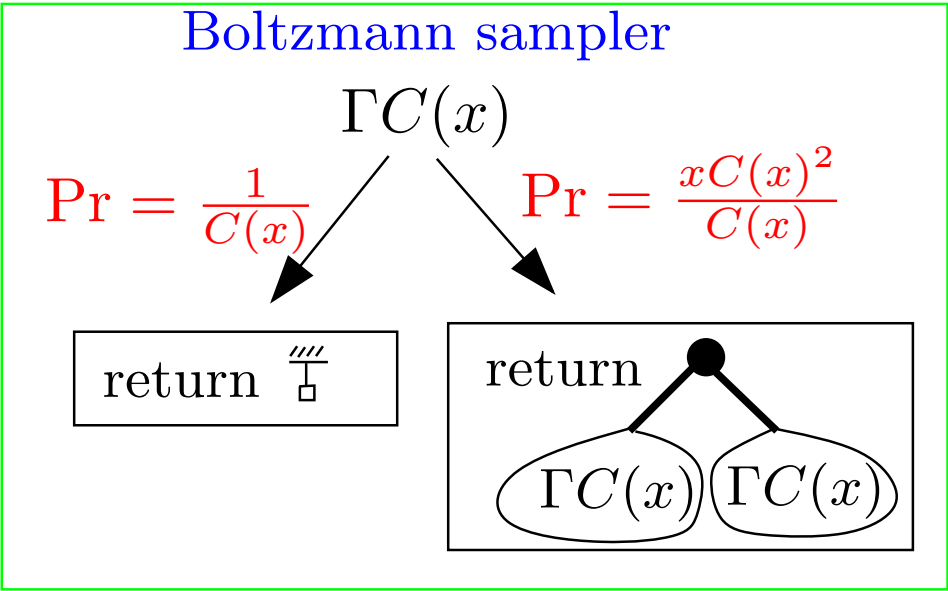
- Based on **gen. funct.**

$$\mathcal{C} = \mathcal{A} + \mathcal{B} \Rightarrow \mathbb{P}(\gamma \in \mathcal{A}) = \frac{A(x)}{C(x)}$$

# Boltzmann samplers: example



Generating function  
 $C(x) = 1 + xC(x)^2$



# Results

**Theorem:** [Duchon et al'02], [Flajolet et al'07]

For any class  $\mathcal{A}$  decomposable in terms of

- basic classes  $\{\mathbf{1}, \mathcal{Z}\}$
- constructions  $\{+, \times, \text{Seq}, \text{Cyc}, \text{Set}\}$ ,

there is a linear-time Boltzmann sampler  $\Gamma_{\tilde{\mathcal{A}}}(x)$ .

**Theorem:** [Bodirsky et al'07,10] For any class  $\mathcal{A}$

decomposable in terms of

- basic classes  $\{\mathbf{1}, \mathcal{Z}, \text{Seq}, \text{Cyc}, \text{Set}\}$ ,
- constructions  $\{+, \times, \circ\}$

there is a linear time Boltzmann sampler  $\Gamma_{\tilde{\mathcal{A}}}(x)$

# Results

**Theorem:** [Duchon et al'02], [Flajolet et al'07]

For any class  $\mathcal{A}$  decomposable in terms of

- basic classes  $\{\mathbf{1}, \mathcal{Z}\}$
- constructions  $\{+, \times, \text{Seq}, \text{Cyc}, \text{Set}\}$ ,

there is a linear-time Boltzmann sampler  $\Gamma\tilde{\mathcal{A}}(x)$ .

**Theorem:** [Bodirsky et al'07,10] For any class  $\mathcal{A}$

decomposable in terms of

- basic classes  $\{\mathbf{1}, \mathcal{Z}, \text{Seq}, \text{Cyc}, \text{Set}\}$ ,
- constructions  $\{+, \times, \circ\}$  and  $\odot$ ,
- cycle-pointing operator  $\mathcal{C} \mapsto \mathcal{C}^\circ$ ,

there is a linear time Boltzmann sampler  $\Gamma\tilde{\mathcal{A}}(x)$  (or  $\Gamma\tilde{\mathcal{A}}^\circ(x)$ ).

# Pólya-Boltzmann samplers

- Ordinary Boltzmann samplers:

$$\tilde{A}(x) = \sum_{\gamma \in \tilde{\mathcal{A}}} x^{|\gamma|} \Rightarrow \mathbb{P}(\gamma) = \frac{x^{|\gamma|}}{\tilde{A}(x)}$$

# Pólya-Boltzmann samplers

- Ordinary Boltzmann samplers:

$$\tilde{A}(x) = \sum_{\gamma \in \tilde{\mathcal{A}}} x^{|\gamma|} \Rightarrow \mathbb{P}(\gamma) = \frac{x^{|\gamma|}}{\tilde{A}(x)}$$

- Pólya-Boltzmann samplers:

$$Z_{\mathcal{A}} = \sum_{\sigma \cdot \gamma = \gamma} W_{(\sigma, \gamma)} \Rightarrow \mathbb{P}(\sigma, \gamma) = \frac{W_{(\sigma, \gamma)}}{Z_{\mathcal{A}}}$$



# Pólya-Boltzmann samplers

- Ordinary Boltzmann samplers:

$$\tilde{A}(x) = \sum_{\gamma \in \tilde{\mathcal{A}}} x^{|\gamma|} \Rightarrow \mathbb{P}(\gamma) = \frac{x^{|\gamma|}}{\tilde{A}(x)}$$

- Pólya-Boltzmann samplers:

$$Z_{\mathcal{A}} = \sum_{\sigma \cdot \gamma = \gamma} W_{(\sigma, \gamma)} \Rightarrow \mathbb{P}(\sigma, \gamma) = \frac{W_{(\sigma, \gamma)}}{Z_{\mathcal{A}}}$$

- Sampling rules  $\{+, \times, \circ\}$  // computation rules for  $Z_{\mathcal{A}}$ :

$$\mathcal{C} = \mathcal{A} + \mathcal{B} \Rightarrow \begin{cases} Z_{\mathcal{C}} = Z_{\mathcal{A}} + Z_{\mathcal{B}} \\ \Gamma Z_{\mathcal{C}} : \text{Bern} \left( \frac{Z_{\mathcal{A}}}{Z_{\mathcal{C}}} \mid \frac{Z_{\mathcal{B}}}{Z_{\mathcal{C}}} \right) \rightarrow \Gamma Z_{\mathcal{A}} \mid \Gamma Z_{\mathcal{B}} \end{cases}$$

# Pólya-Boltzmann samplers

- Ordinary Boltzmann samplers:

$$\tilde{A}(x) = \sum_{\gamma \in \tilde{\mathcal{A}}} x^{|\gamma|} \Rightarrow \mathbb{P}(\gamma) = \frac{x^{|\gamma|}}{\tilde{A}(x)}$$

- Pólya-Boltzmann samplers:

$$Z_{\mathcal{A}} = \sum_{\sigma \cdot \gamma = \gamma} W_{(\sigma, \gamma)} \Rightarrow \mathbb{P}(\sigma, \gamma) = \frac{W_{(\sigma, \gamma)}}{Z_{\mathcal{A}}}$$

- Sampling rules  $\{+, \times, \circ\}$  // computation rules for  $Z_{\mathcal{A}}$ :

$$\mathcal{C} = \mathcal{A} + \mathcal{B} \Rightarrow \begin{cases} Z_{\mathcal{C}} = Z_{\mathcal{A}} + Z_{\mathcal{B}} \\ \Gamma Z_{\mathcal{C}} : \text{Bern} \left( \frac{Z_{\mathcal{A}}}{Z_{\mathcal{C}}} \mid \frac{Z_{\mathcal{B}}}{Z_{\mathcal{C}}} \right) \rightarrow \Gamma Z_{\mathcal{A}} \mid \Gamma Z_{\mathcal{B}} \end{cases}$$

- Recover ordinary Boltzmann sampler using specialization

$$Z_{\mathcal{C}}(x, x^2, \dots) = C(x) \Rightarrow \Gamma Z_{\mathcal{C}}(x, x^2, \dots) = \Gamma \tilde{C}(x)$$

# Sampler for trees

Let  $t(x)$  be the OGF of (unrooted) trees.

1) Translate the equation

$$R(x) = \exp\left(\sum_{i \geq 1} \frac{1}{i} R(x^i)\right)$$

into a **Boltzmann sampler for  $\mathcal{R}$**  cf [Flajolet et al'07]  
(superposition of Poisson laws)

2) Translate the equation

$$xt'(x) = \underbrace{R(x)}_{\text{rooted}} + \underbrace{x^2 R'(x^2)}_{\substack{\text{centre symmetry} \\ \text{is edge}}} + \underbrace{\sum_{i \geq 2} x^i R'(x^i) R(x)}_{\substack{\text{centre symmetry} \\ \text{is vertex}}}$$

into a **Boltzmann sampler for  $\widetilde{\mathcal{T}}^\circ$** .

# Sampler for trees

Let  $t(x)$  be the OGF of (unrooted) trees.

1) Translate the equation

$$R(x) = \exp\left(\sum_{i \geq 1} \frac{1}{i} R(x^i)\right)$$

into a **Boltzmann sampler for  $\mathcal{R}$**  cf [Flajolet et al'07]  
(superposition of Poisson laws)

2) Translate the equation

$$xt'(x) = \underbrace{R(x)}_{\text{rooted}} + \underbrace{x^2 R'(x^2)}_{\substack{\text{centre symmetry} \\ \text{is edge}}} + \underbrace{\sum_{i \geq 2} x^i R'(x^i) R(x)}_{\substack{\text{centre symmetry} \\ \text{is vertex}}}$$

into a **Boltzmann sampler for  $\widetilde{\mathcal{T}}^\circ$** .

(Also **recursive sampler** by Wilf using “centre of gravity”)

# Open problems

- Refined **complexity analysis** of Pólya-Boltzmann samplers, cf [Pivoteau, Salvy, Soria'08]
- Boltzmann sampling with **catalytic variables**
- For which recursive specification can we **determine the growth rate** automatically ?
- Use Boltzmann samplers to **study random unlabelled structures**, cf [Panagiotou, Steger].