A linear time algorithm for the random generation of labeled planar graphs

Éric Fusy

Algorithms Project, INRIA Rocquencourt

Plan

- Principles of Boltzmann samplers.

- Application to planar graphs

- Size distribution and complexity results:
  - A linear time approximate size random generator of planar graphs
  - A quadratic time exact size random generator of planar graphs

- Implementation and experimentations

The general framework of Boltzmann samplers

Idea of Boltzmann samplers

- Introduced by Duchon, Flajolet, Louchard and Schaeffer (2002)
- Relax the constraint of fixed size (cf recursive method) for random generation.
- The distribution is spread over all objects of the class.
- An object is drawn with probability proportional to the exponential of its size (cf statistical physics)

Unlabelled sets

- Let $\mathcal{C}$ be an unlabelled combinatorial class (e.g. binary trees)
  Ordinary generating function:

$$C(x) = \sum_{\gamma \in \mathcal{C}} x^{|\gamma|} = \sum_{n \geq 0} c_n x^n,$$

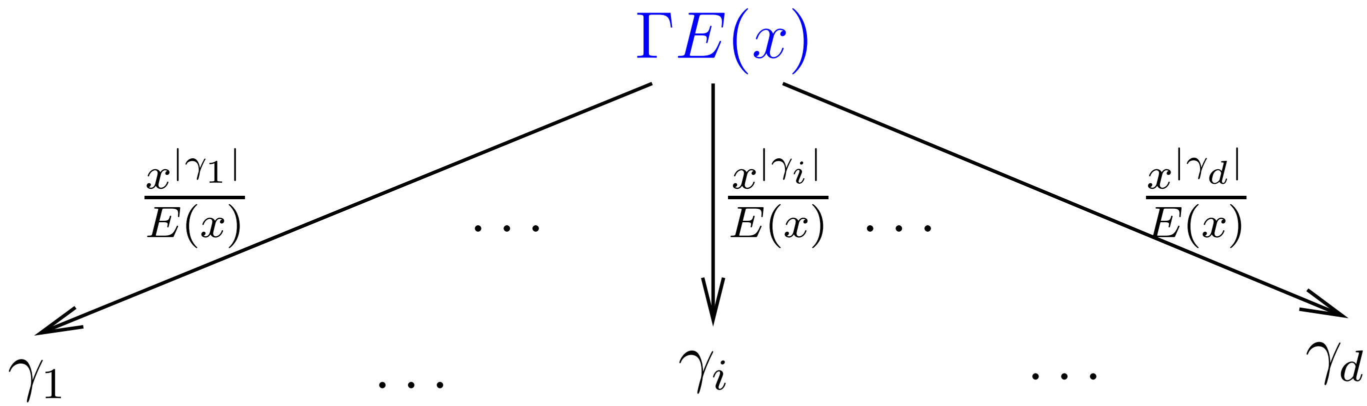
  where $|\gamma|$ is the size of γ.

- Given $x > 0$ $(x \leq \rho_C)$ a fixed real value,
  a Boltzmann sampler $\Gamma C(x)$ is a procedure that draws each object γ of $\mathcal{C}$ with probability:

$$Pr(\gamma) = \frac{x^{|\gamma|}}{C(x)}$$

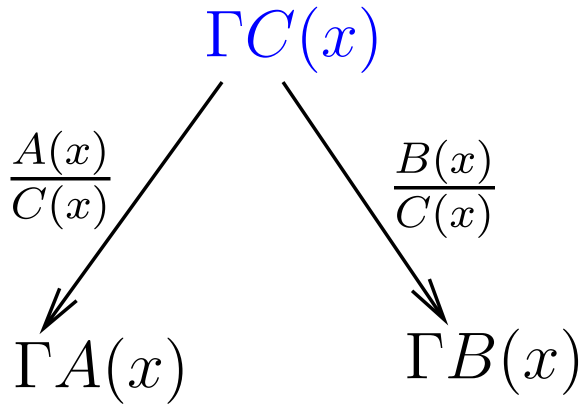
Finite sets

Let $\mathcal{E} = (\gamma_1, \ldots, \gamma_d)$ $\quad\quad$ $E(x) = \sum_{i=1}^{d} x^{|\gamma_i|}$



The basic construction rules

Union: Let $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$. Assume we have Boltzmann samplers $\Gamma A(x)$ for $\mathcal{A}$ and $\Gamma B(x)$ for $\mathcal{B}$. Define $\Gamma C(x)$ as:



$\Rightarrow$ $\Gamma C(x)$ is a Boltzmann sampler for $\mathcal{A} \cup \mathcal{B}$.

Proof:

- If $\gamma \in \mathcal{A}$, then $Pr(\gamma) = \frac{A(x)}{C(x)} \cdot \frac{x^{|\gamma|}}{A(x)} = \frac{x^{|\gamma|}}{C(x)}$.
- If $\gamma \in \mathcal{B}$, then $Pr(\gamma) = \frac{B(x)}{C(x)} \cdot \frac{x^{|\gamma|}}{B(x)} = \frac{x^{|\gamma|}}{C(x)}$.

The basic construction rules

Product: Let $\mathcal{C} = \mathcal{A} \times \mathcal{B}$. Assume we have Boltzmann samplers $\Gamma A(x)$ for $\mathcal{A}$ and $\Gamma B(x)$ for $\mathcal{B}$. Define $\Gamma C(x)$ as:

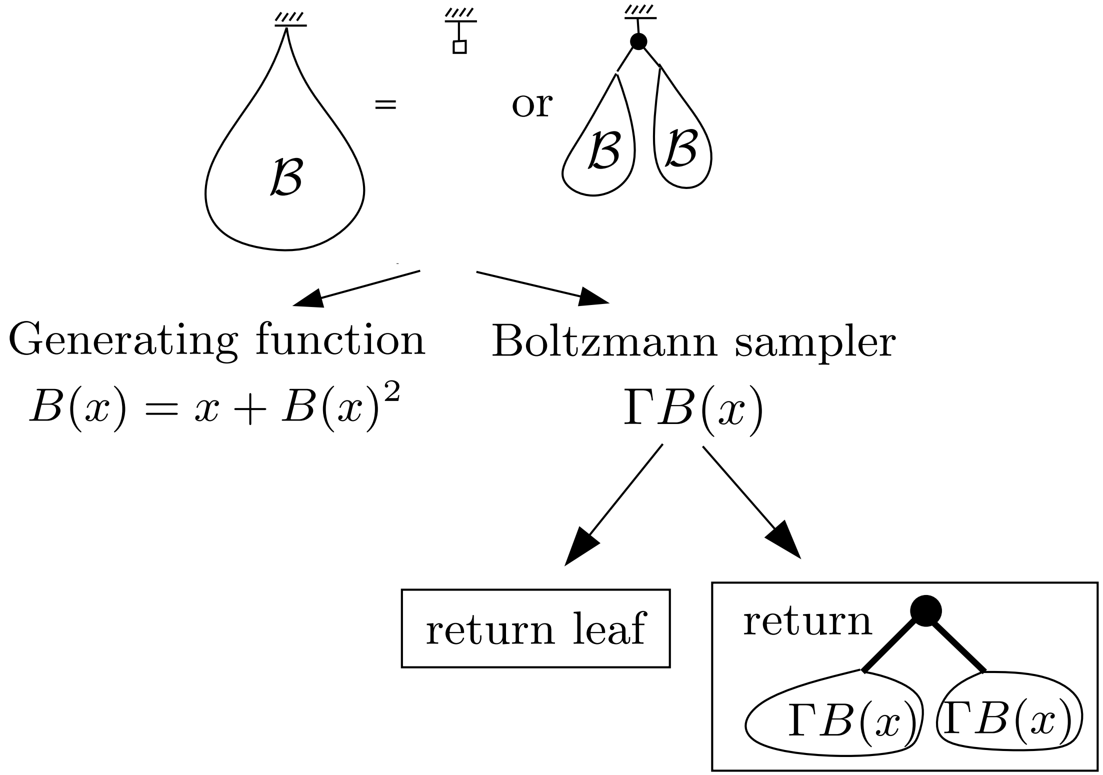
$$\Gamma C(x) \quad : \quad \begin{array}{l} \gamma_1 \leftarrow \Gamma A(x) \\ \gamma_2 \leftarrow \Gamma B(x) \\ \textsf{return } (\gamma_1, \gamma_2) \end{array}$$

$\Rightarrow$ $\Gamma C(x)$ is a Boltzmann sampler for $\mathcal{A} \cup \mathcal{B}$:

Proof: an object $\gamma = (\gamma_1, \gamma_2)$ has probability:

$$\frac{x^{|\gamma_1|}}{A(x)} \frac{x^{|\gamma_2|}}{B(x)} = \frac{x^{|\gamma_1|+|\gamma_2|}}{A(x) \cdot B(x)} = \frac{x^{|\gamma|}}{C(x)}$$

Example: binary trees



Result for unlabeled sets

Theorem:

- A Boltzmann sampler can be assembled for an unlabeled class specified with the constructions $\cup$, $\times$, $\mathrm{Sequence}$.
- The complexity is linear in the size of the output object.

Construction	Boltzmann sampler $\Gamma C(x)$
$\mathcal{C} = \varnothing$	return $\varnothing$
$\mathcal{C} = \bullet$	return $\bullet$
$\mathcal{C} = \mathcal{A} \cup \mathcal{B}$	$\mathrm{Bern}\left(\frac{A(x)}{C(x)} \middle| \frac{B(x)}{C(x)}\right)$? $\Gamma A(x) \vert \Gamma B(x)$
$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	return $(\Gamma A(x), \Gamma B(x))$
$\mathcal{C} = \mathrm{Seq}\,(\mathcal{A})$	$k \leftarrow \mathrm{Geom}\,(A(x))$ return $(\Gamma A(x), \ldots, \Gamma A(x))$ { k calls}

Labeled sets

- Let $\mathcal{C}$ be a labeled combinatorial class (e.g. permutations)
  Exponential generating function:

$$C(x) = \sum_{\gamma \in \mathcal{C}} \frac{x^{|\gamma|}}{|\gamma|!} = \sum_n c_n \frac{x^n}{n!},$$

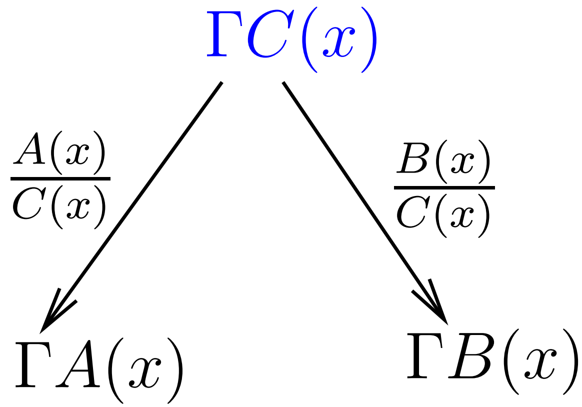
  where $|\gamma|$ is the size of γ.

- Given $x > 0$ $(x \leq \rho_C)$ a fixed real value, a Boltzmann sampler $\Gamma C(x)$ draws each object γ of $\mathcal{C}$ with probability:

$$Pr(\gamma) = \frac{1}{C(x)} \frac{x^{|\gamma|}}{|\gamma|!}$$

The basic construction rules

Union: Let $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$. Assume we have Boltzmann samplers $\Gamma A(x)$ for $\mathcal{A}$ and $\Gamma B(x)$ for $\mathcal{B}$. Define $\Gamma C(x)$ as:



$\Rightarrow$ $\Gamma C(x)$ is a Boltzmann sampler for $\mathcal{A} \cup \mathcal{B}$:

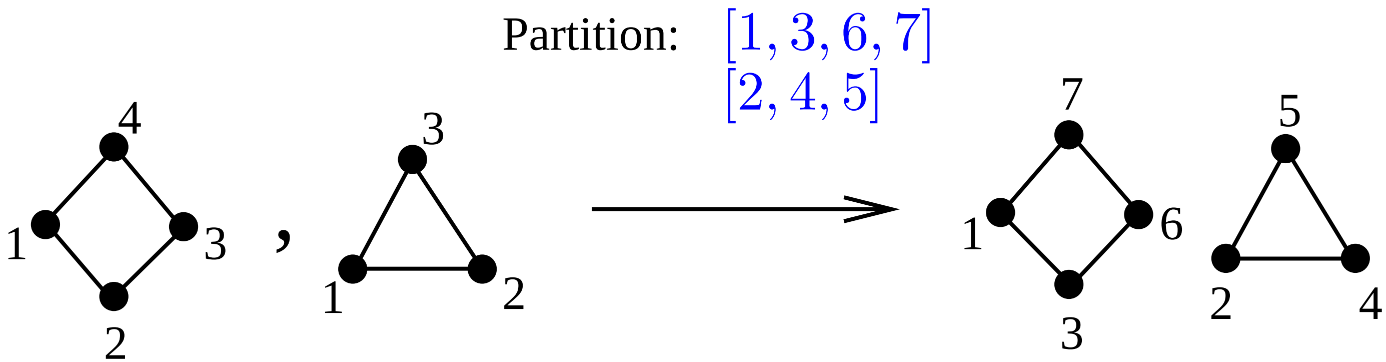
Proof:

- If $\gamma \in \mathcal{A}$, then $Pr(\gamma) = \frac{A(x)}{C(x)} \cdot \left(\frac{1}{A(x)} \frac{x^{|\gamma|}}{|\gamma|!} \right) = \frac{1}{C(x)} \frac{x^{|\gamma|}}{|\gamma|!}$.
- If $\gamma \in \mathcal{B}$, then $Pr(\gamma) = \frac{B(x)}{C(x)} \cdot \left(\frac{1}{C(x)} \frac{x^{|\gamma|}}{|\gamma|!} \right) = \frac{1}{B(x)} \frac{x^{|\gamma|}}{|\gamma|!}$.

Cartesian product for labelled sets

An object of $\mathcal{A} \star \mathcal{B}$ is obtained by:

- taking a pair (γ_1, γ_2) with $\gamma_1 \in \mathcal{A}$ and $\gamma_2 \in \mathcal{B}$.
- Relabel according to a partition of $[1, \ldots, |\gamma_1| + |\gamma_2|]$.

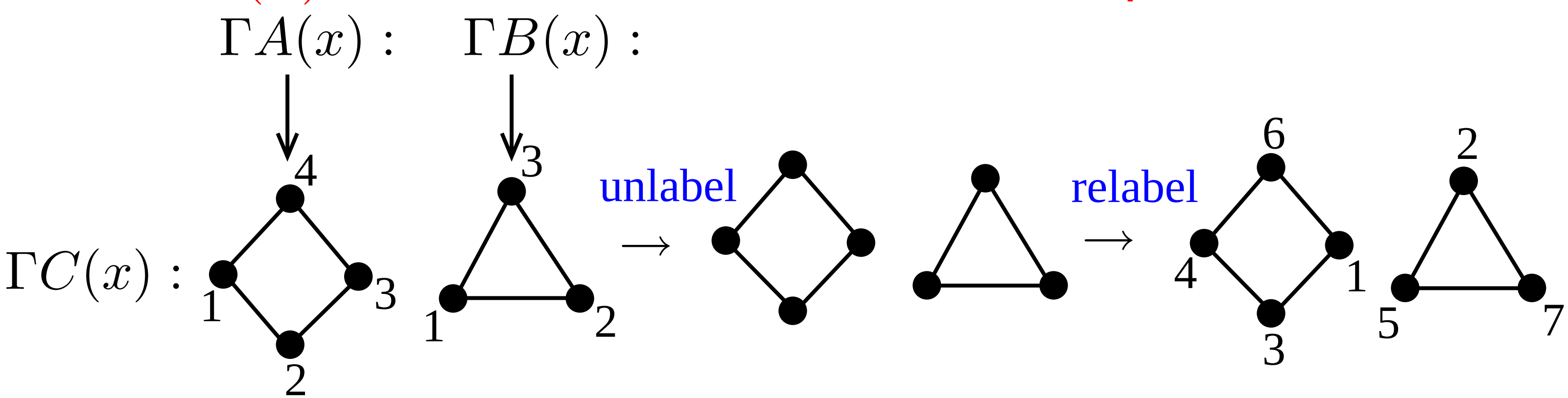


Boltzmann for cartesian product

Cartesian product: Let $\mathcal{C} = \mathcal{A} \star \mathcal{B}$. Assume we have Boltzmann samplers $\Gamma A(x)$ for $\mathcal{A}$ and $\Gamma B(x)$ for $\mathcal{B}$. Define $\Gamma C(x)$ as:

$\Gamma C(x)$: $\gamma_1 \leftarrow \Gamma A(x)$
$\gamma_2 \leftarrow \Gamma B(x)$
remove the labels on γ_1 and γ_2
throw disctinct labels at random on (γ_1, γ_2)
return (γ_1, γ_2)

$\Rightarrow$ $\Gamma C(x)$ is a Boltzmann sampler for $\mathcal{A} \star \mathcal{B}$



Result for labeled sets

Theorem:

- A Boltzmann sampler can be assembled for a labeled class specified with the constructions $\cup$, $\times$, Set.
- The complexity is linear in the size of the output object.
- The labels have just to be thrown at the end.

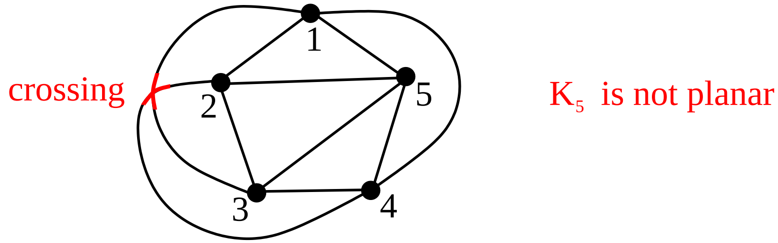
Boltzmann vs the recursive method

	Boltzmann	recursive method
size distribution	$Pr(size = n) = \frac{C_n x^n}{C(x)}$	fixed size n
auxiliary memory	$\mathcal{O}(\log(n))$	$\mathcal{O}(n^2)$
time per generation	$\mathcal{O}(n^2)$ Exact $\mathcal{O}(n)$ Approx	$\mathcal{O}(n \log(n))$ Exact

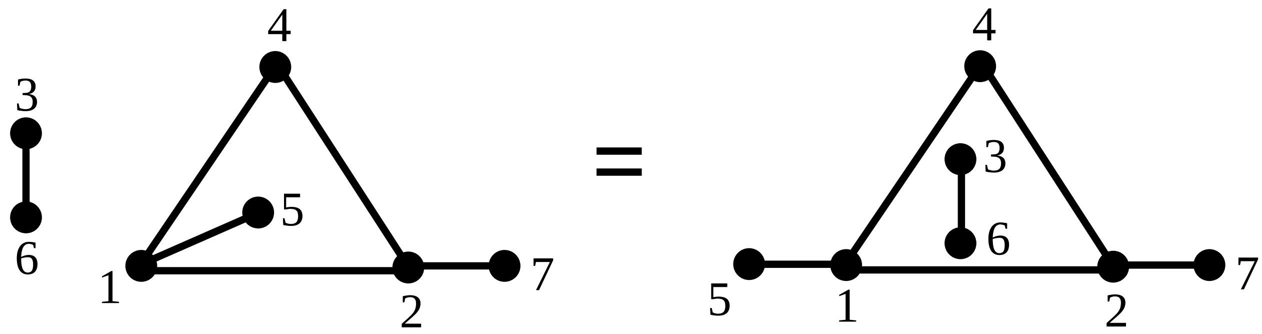
Planar graphs and Boltzmann samplers

Labeled Planar graphs

- A labelled graph with n vertices is a set of edges on the labeled vertex-set $V = [1, \ldots, n]$.
- A graph is planar if it can be embedded in the plane.



- The embedding does not count ($\neq$ planar maps)



Random generation of planar graphs

Existing algorithms:

- Markov chain (Denise, Vasconcellos, Welsh): simple algorithm but unknown convergence rate (mixing time)
- Recursive method (Bodirsky, Gröpl, Kang): Polynomial time algorithm for uniform random generation of planar graphs with n vertices but large preprocessing time (many coefficients need to be stored).

What new has to be done?

To design a Boltzmann sampler for labeled planar graphs, we have to do the following:

- A planar graph has labeled vertices and unlabeled edges: $\Rightarrow$ define the Boltzmann framework for the case of a mixed class (two variables)
- Add the substitution (composition of G.f.) to the constructions.
- Add rejection techniques to do derooting/rerooting operations on the graphs

Boltzmann samplers: mixed classes

- Let $\mathcal{C}$ be a mixed combinatorial class (e.g. planar graphs)
  Mixed Generating function:

$$C(x,y) = \sum_{\gamma \in \mathcal{C}} \frac{x^{i(\gamma)}}{i(\gamma)!} y^{j(\gamma)} = \sum_{i,j} c_{i,j} \frac{x^i}{i!} y^j,$$

  $i(\gamma)$ is the number of labeled atoms (e.g. vertices)
  $j(\gamma)$ is the number of unlabeled atoms (e.g. edges)

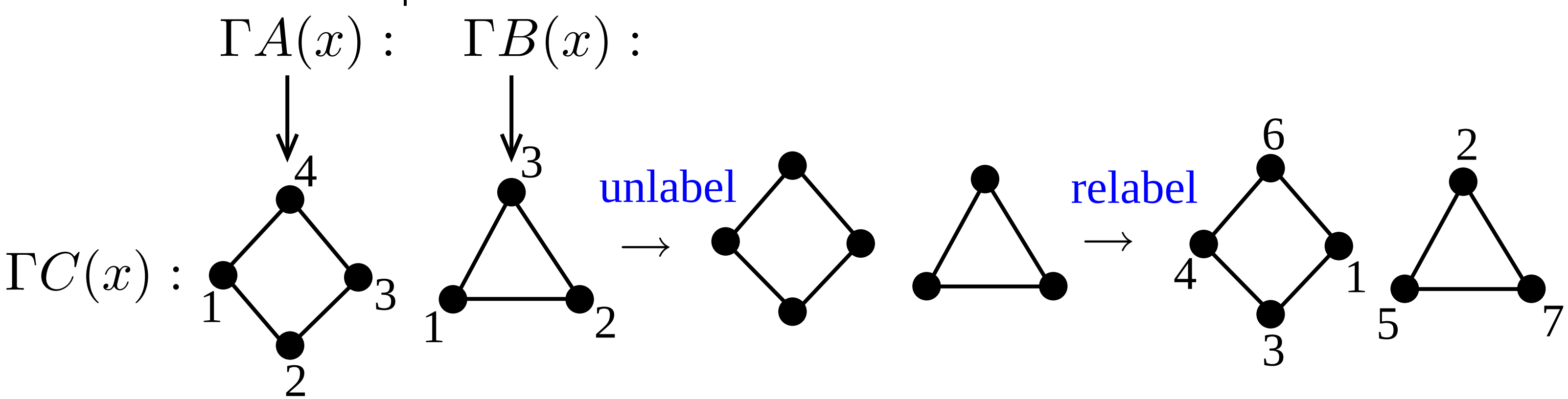
- Given $x > 0$ and $y > 0$ two fixed real values, a Boltzmann sampler $\Gamma C(x,y)$ draws each object γ of $\mathcal{C}$ with probability:

$$Pr(\gamma) = \frac{1}{C(x,y)} \frac{x^{i(\gamma)}}{i(\gamma)!} y^{j(\gamma)}$$

Boltzmann samplers: mixed classes

Theorem: A Boltzmann sampler can be assembled for a mixed class specified with the constructions $\cup$, $\times$, Set.
Linear time complexity in the size of the output object.

Construction	Boltzmann sampler $\Gamma C(x,y)$
$\mathcal{C} = \mathcal{A} \cup \mathcal{B}$	$\mathrm{Bern}\left(\frac{A(x,y)}{C(x,y)}\right)?\ \Gamma A(x,y) \vert \Gamma B(x,y)$
$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	return $(\Gamma A(x,y), \Gamma B(x,y))$+relabel
$\mathcal{C} = \mathrm{Set}\,(\mathcal{A})$	$\mathrm{nr\_components} \leftarrow \mathrm{Poiss}\,(A(x,y))$ return $(\Gamma A(x,y), \ldots, \Gamma A(x,y))$ +relabel



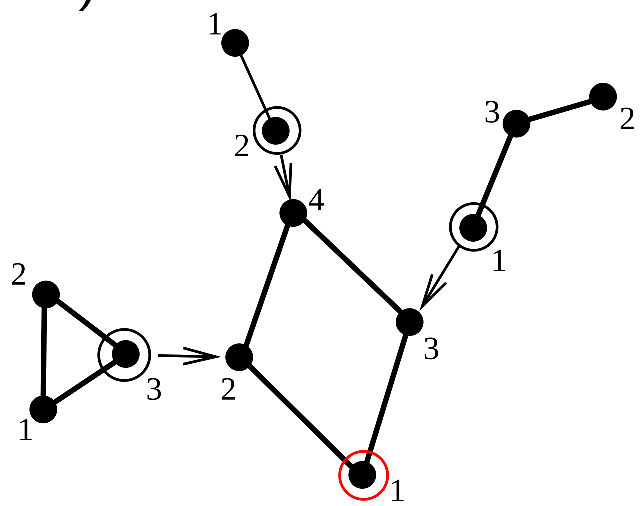
Boltzmann samplers: substitution

- The class $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ consists of objects of $\mathcal{A}$ where each atom is replaced by an object of $\mathcal{B}$
  G.f.: $C(x) = A(B(x))$
- Boltzmann sampler:

$$\Gamma C(x) \quad \gamma \leftarrow \Gamma A\,(B(x))$$

replace each atom of γ by $\Gamma B(x)$

- very simple and no need of Bernoulli-choices (unlike the recursive method)



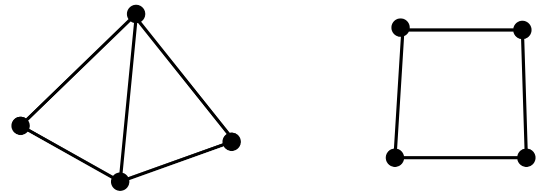
Conception of a Boltzmann sampler for planar graphs

Overview of the method

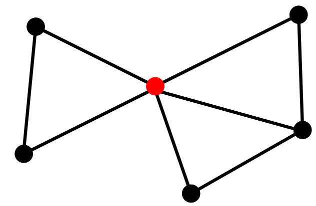
- Decomposition according to successive levels of connectivity:
  Planar graph $\longrightarrow$ Connected $\longrightarrow$ 2-connected $\longrightarrow$ 3-connected
- Combinatorial bijection (Fusy, Poulalhon, Schaeffer)
  3-connected graphs $\leftrightarrow$ binary trees

Planar graphs

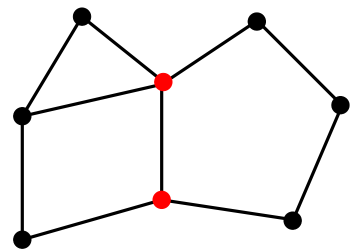
Disconnected



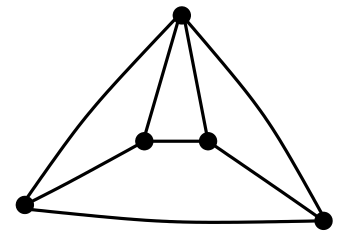
Connected



2-connected



3-connected



Planar graphs $\rightarrow$ connected p. g.

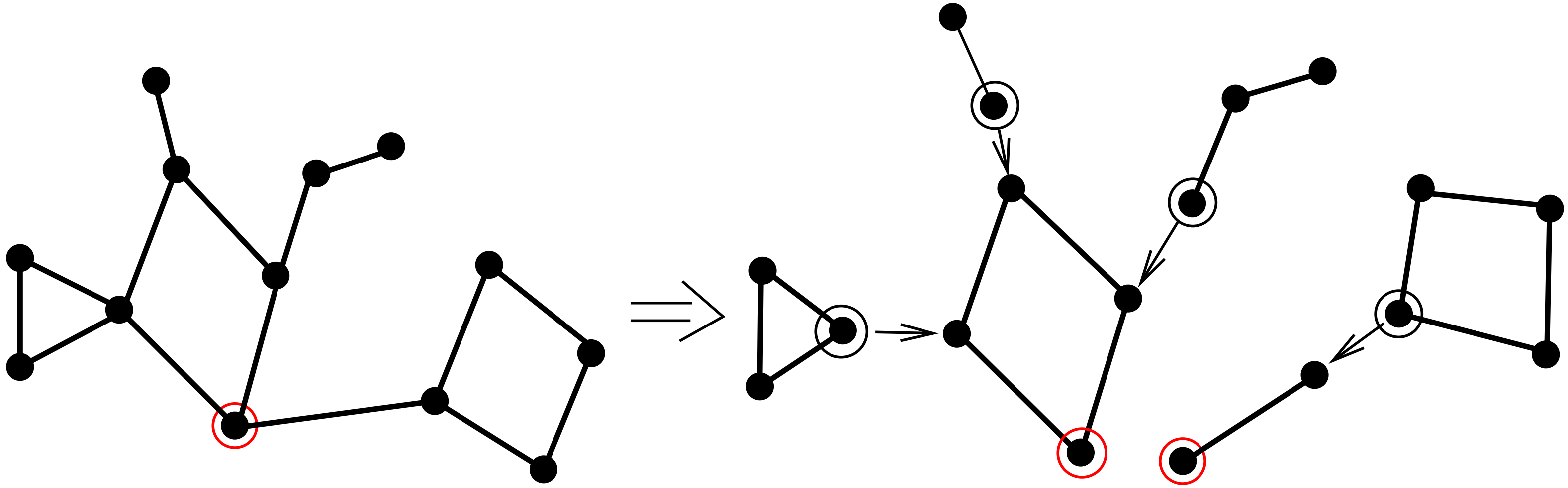
- Let $\mathcal{G}$ be the set of planar graphs
- Let $\mathcal{C}$ be the set of connected planar graphs
- A planar graph is decomposed into connected components
  $\Rightarrow \mathcal{G} = Set(\mathcal{C})$ $\quad\quad G(x) = \exp(C(x))$

$$\Gamma G(x,y) \quad : \quad k \leftarrow Poiss(C(x,y))$$
$$\text{return } (\Gamma C(x,y), \ldots, \Gamma C(x,y)) \ \{ \ k \text{ calls}\}$$

Connected → 2-connected

Decomposition by vertex-substitution:
A pointed connected planar graph is a set of pointed 2-connected planar graphs where each non pointed vertex is substituted by a pointed connected planar graph.

$\Rightarrow$ $\boxed{C^\bullet(x,y) = x\exp(B'(C^\bullet(x,y),y))}$



Connected $\longrightarrow$ 2-connected

$$C^\bullet(x) = x\exp(B'(C^\bullet(x)))$$

$\Gamma C^\bullet(x)$: 1) $k \leftarrow Poiss(\lambda := B'(C^\bullet(x))$ $\qquad \exp(\ldots)$

2) $\gamma \leftarrow \underbrace{(\Gamma B^\bullet(C^\bullet(x)), \ldots, \Gamma B^\bullet(C^\bullet(x)))}_{k\ times}$ $\qquad \exp(B'(\ldots))$

3) merge the k marked vertices of γ

4) for each non-marked vertex v of γ

substitute v by $\gamma_v \leftarrow \Gamma C^\bullet(x)$ $\qquad \exp(B'(C^\bullet(x)))$

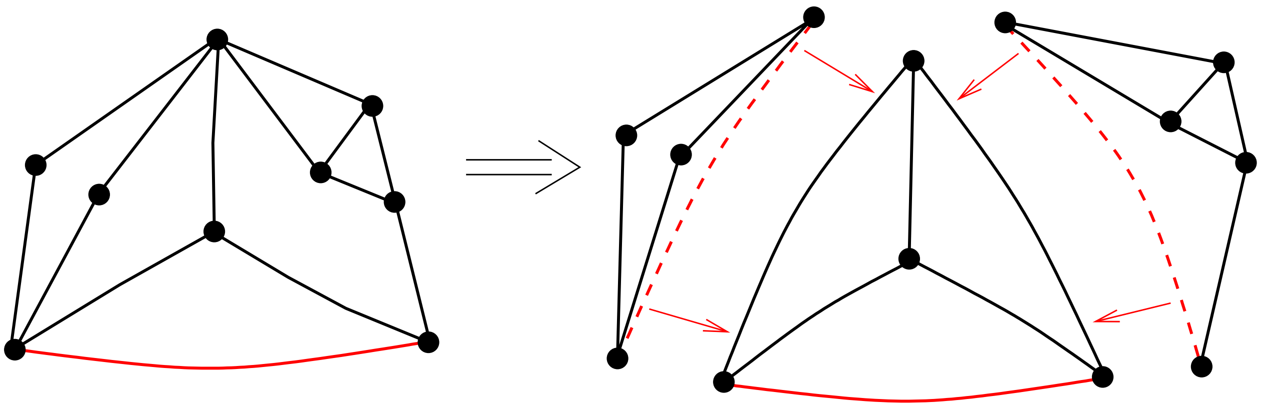
5) return γ

$\Rightarrow$ Finding $\Gamma C^\bullet$ reduces to finding $\Gamma B^\bullet$

2-connected $\rightarrow$ 3-connected

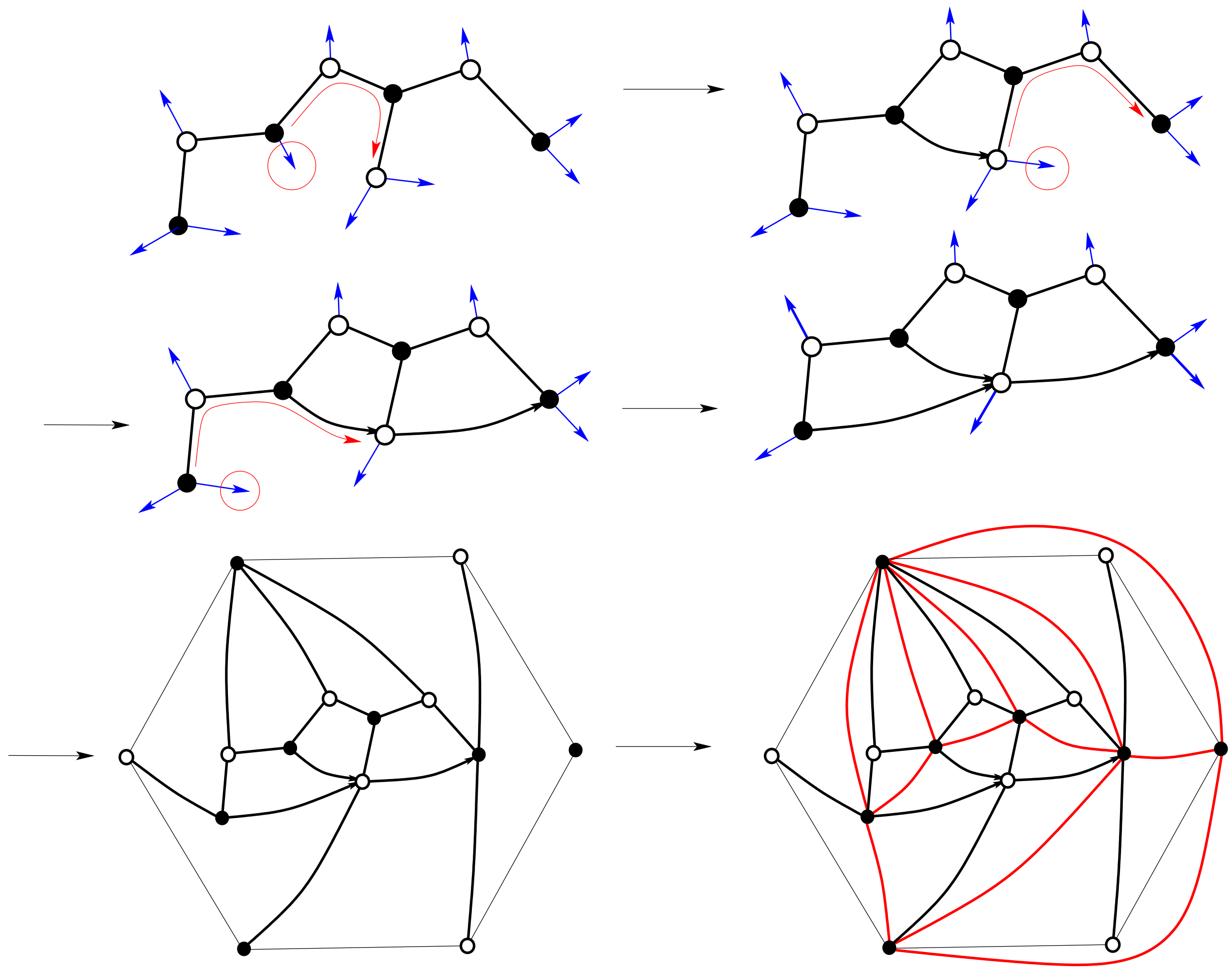
- Decomposition by edge-substitution.
- $B(x,y)$ series of 2-connected planar graphs.
- $G_3(x,y)$ series of 3-connected planar graphs

$$\frac{\partial B}{\partial y}(x,y) \approx \frac{\partial G_3}{\partial y}\left(x, \frac{\partial B}{\partial y}(x,y)\right)$$



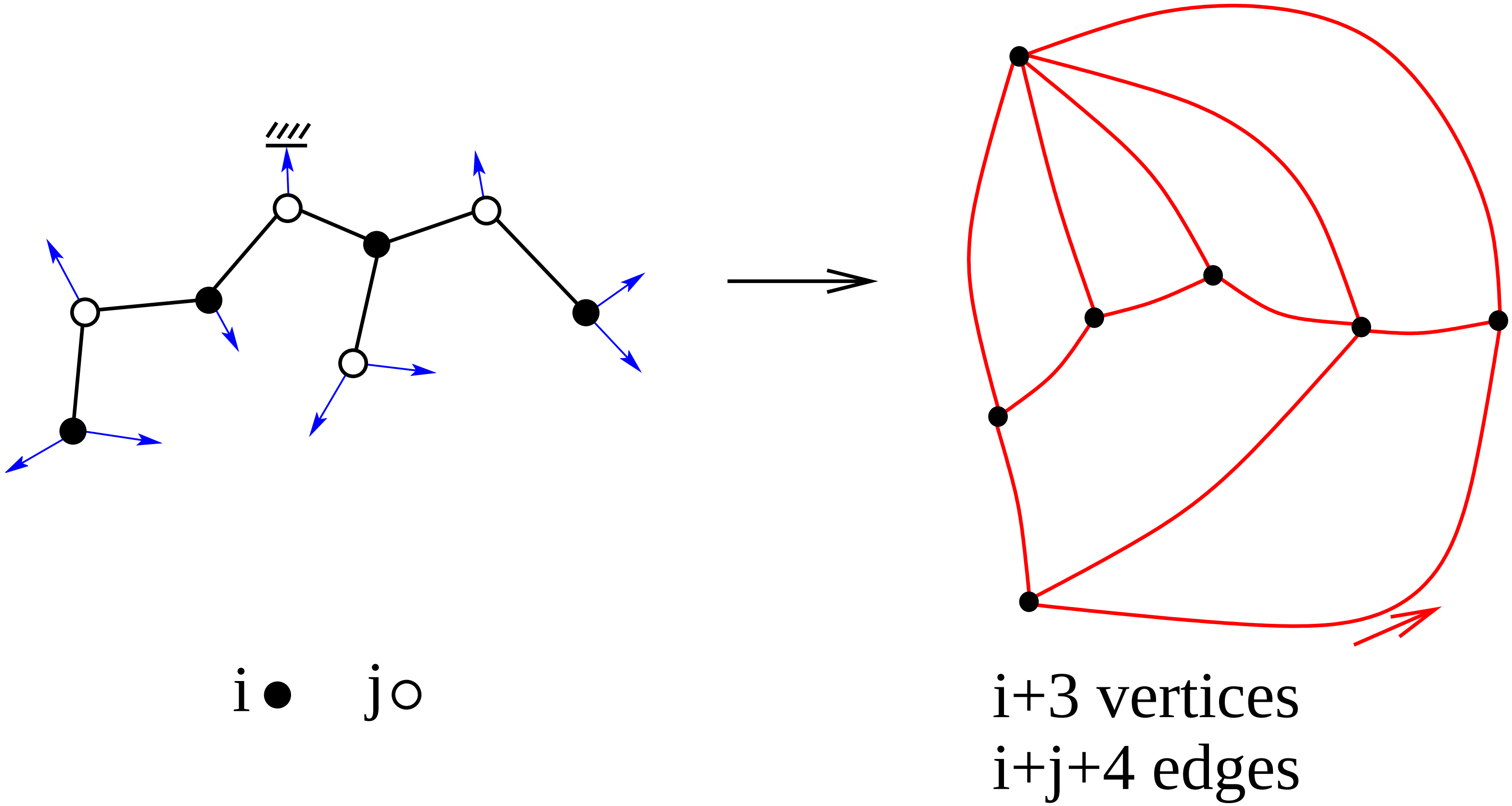
$\Rightarrow$ Finding $\Gamma\frac{\partial B}{\partial y}$ reduces to finding $\Gamma\frac{\partial G_3}{\partial y}$

3-connected $\leftrightarrow$ binary trees

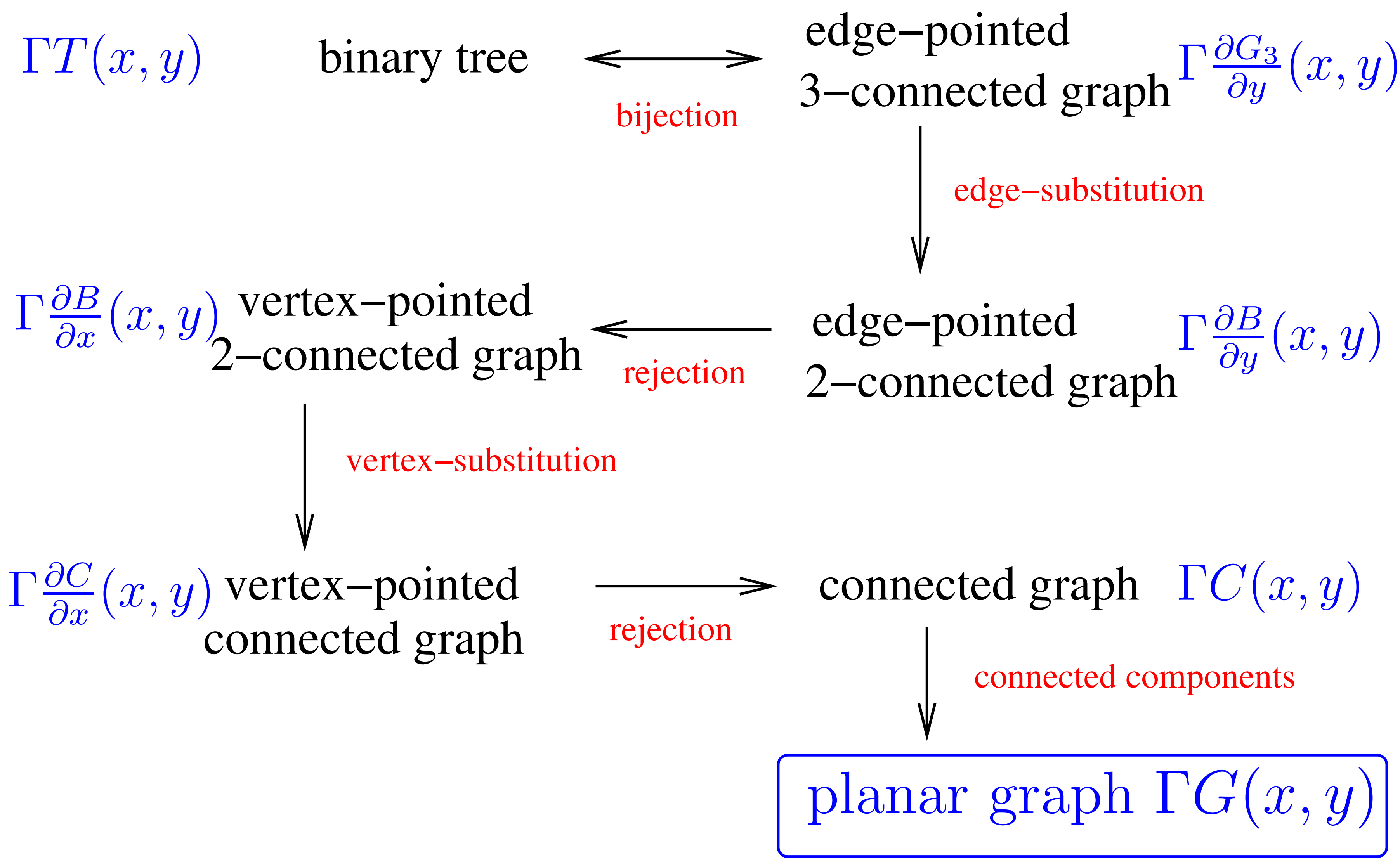


3-connected ↔ binary trees

Fusy, Poulalhon, Schaeffer 2005:
Binary trees are in bijection with edge-pointed 3-connected planar graphs.



Boltzmann sampler for planar graphs



Rejection for Boltzmann samplers

- Let $\mathcal{B}$ be a combinatorial class for which we have a Boltzmann sampler.
- Let $\mathcal{A} \subset \mathcal{B}$ be a combinatorial class for which we want a Boltzmann sampler.

$$\Gamma A(x) \quad : \quad \gamma \leftarrow \Gamma B(x)$$
$$\text{if } \gamma \in \mathcal{A} \text{ return } \gamma \text{ else restart}$$

Then $\Gamma A(x)$ is a Boltzmann sampler for $\mathcal{A}$.

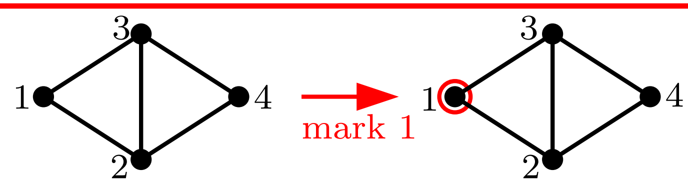
The acceptance probability at each try is

$$P_{accept} = \sum_{\gamma \in \mathcal{A}} \frac{x^{|\gamma|}}{B(x)} = \frac{A(x)}{B(x)}$$

Applications

- $\Gamma C(x,y)$ from $\Gamma C^{\bullet}(x,y)$

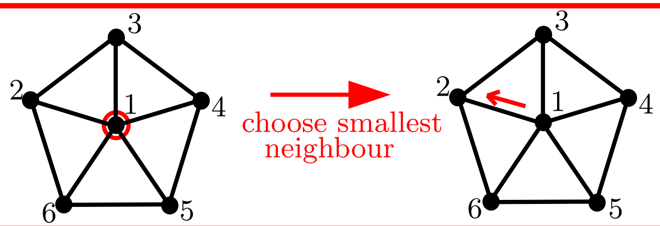
$\mathcal{C}$ is a subset of $\mathcal{C}^{\bullet}$:



$\Rightarrow \Gamma C(x,y):\ \gamma \leftarrow \Gamma C^{\bullet}(x,y)$
if the pointed vertex has label 1, return γ
else restart

- $\Gamma \frac{\partial B}{\partial x}(x,y)$ from $\Gamma \frac{\partial B}{\partial y}(x,y)$

$\mathcal{B}^{\bullet}$ is a subset of $\mathcal{B}^{\rightarrow}$:



$\Rightarrow \Gamma \frac{\partial B}{\partial x}(x,y):\ \gamma \leftarrow \Gamma \frac{\partial B}{\partial y}(x,y)$
if the end of the root-edge is the smallest neighbour of the origin of the root-edge, return γ
else restart

Derivation of an efficient sampler

How to achieve a target size n?

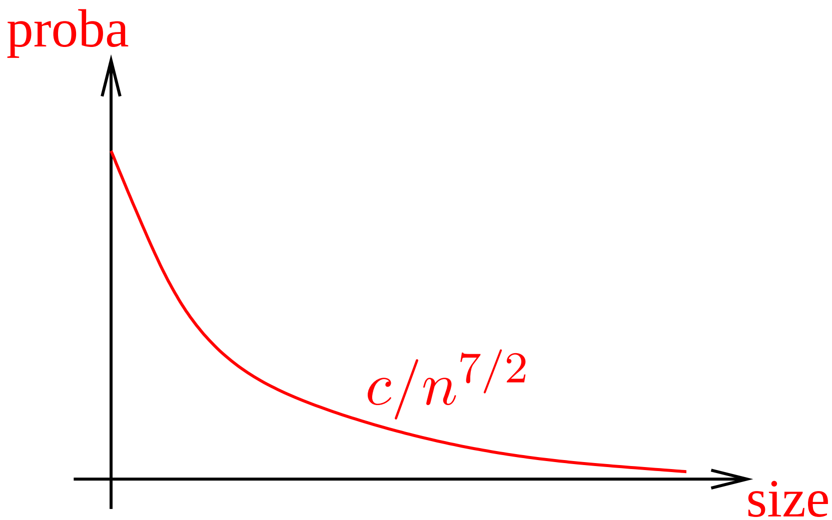
- We have a Boltzmann sampler $\Gamma G(x, 1)$ for planar graphs
- We want to achieve a target-size n
- We have to choose $x = x_n$ so that $\Gamma G(x_n)$ produces graphs of size n with good probability.
- Natural choice: x_n such that $\boxed{\mathbf{E}(size(\Gamma G(x_n))) = n}$
- The function $x \to \mathbf{E}(size(\Gamma G(x)))$ is increasing
$\Rightarrow$ x_n has to converge to ρ_G (dom. sing.) when $n \to \infty$.

Size distribution

Problem: Even at the singularity ρ_G, the expected size of $\Gamma G(\rho_G)$ remains bounded:

$$Pr(size = n) = \frac{G_n \rho_G^n}{G(x)} \sim \frac{c}{n^{7/2}}$$

(Giménez, Noy 2005)

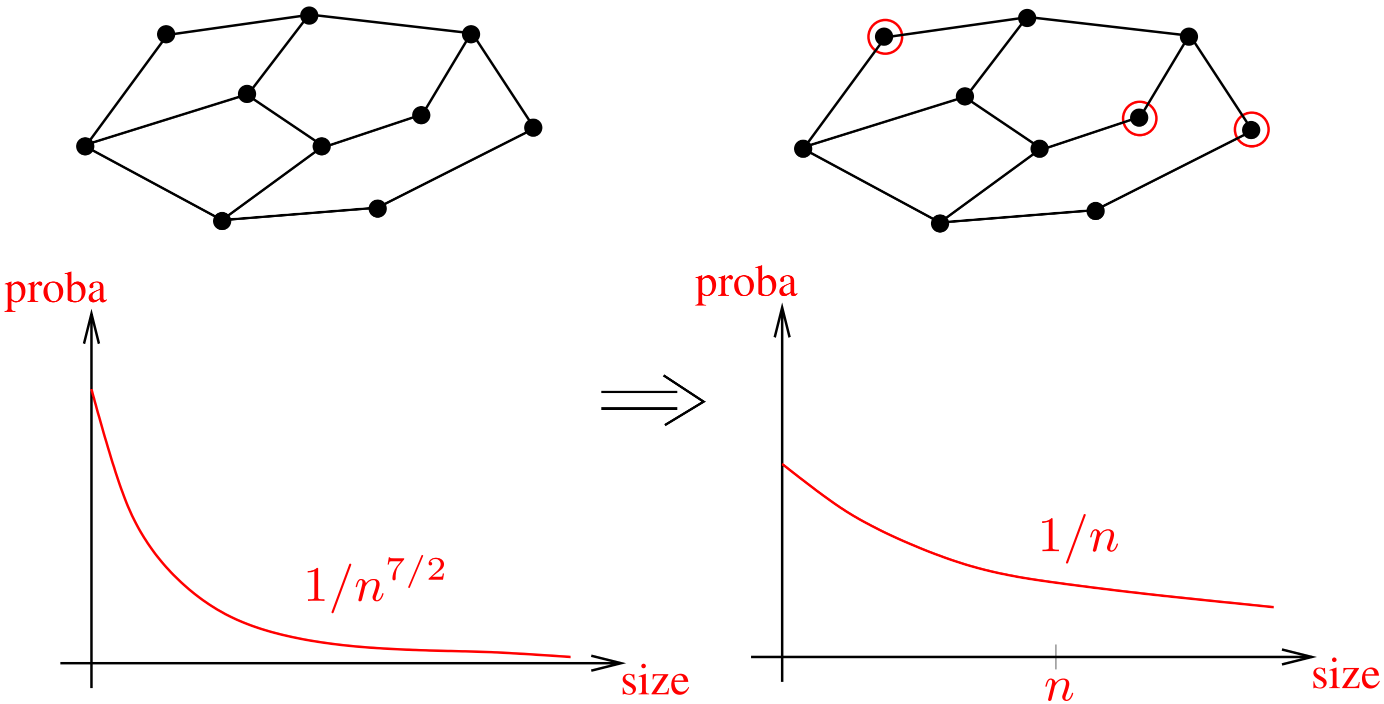


Size distribution of the output of $\Gamma G(\rho_G)$

Improve the size distribution

Solution: point the graphs 3 times

Effect: multiply coefficient G_n by n^3.



Output of $\Gamma G(\rho_G)$ Output of $\Gamma G^{\bullet\bullet\bullet}\left(\rho_G(1-\frac{1}{2n})\right)$

Inject pointing into decomposition

- $\mathcal{C} = \mathcal{A} \cup \mathcal{B} \Rightarrow \mathcal{C}^\bullet = \mathcal{A}^\bullet \cup \mathcal{B}^\bullet$
- $\mathcal{C} = \mathcal{A} \star \mathcal{B} \Rightarrow \mathcal{C}^\bullet = \mathcal{A}^\bullet \star \mathcal{B} \cup \mathcal{A} \star \mathcal{B}^\bullet$
- $\mathcal{C} = Set(\mathcal{A}) \Rightarrow \mathcal{C}^\bullet = \mathcal{A}^\bullet \star Set(\mathcal{A})$

Example: pointed binary trees

$$\begin{cases} B(x) &= x + B(x)^2 \\ B^\bullet(x) &= x + B^\bullet(x)B(x) + B(x)B^\bullet(x) \end{cases}$$

Main results

Let n be a target size and ϵ be a (relative) size-tolerance.
Take $\Gamma G^{\bullet\bullet\bullet}(x_n)$ at $x_n = \rho_G\left(1 - \frac{1}{2n}\right)$.

Theorem The generator $\Gamma G^{\bullet\bullet\bullet}(x_n)$ produces planar graphs:

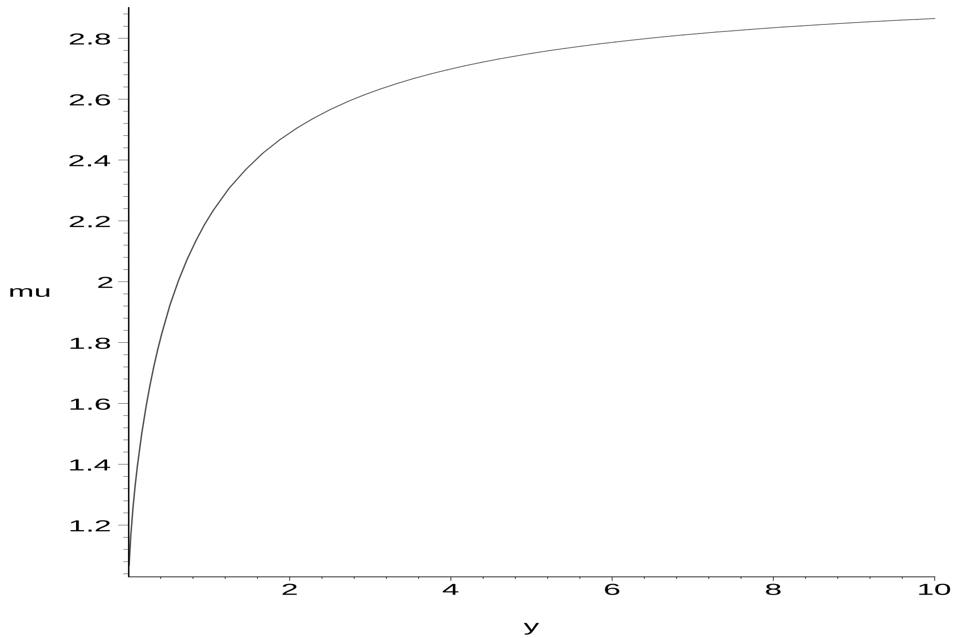
- with size in $[n(1-\epsilon), n(1+\epsilon)]$ in linear time. APPROX
- with size n in quadratic time. EXACT

	Aux. memory	Prep. time	Time per generation
Markov	$\mathcal{O}(\log n)$	$\mathcal{O}(1)$	*unknown* {exact size}
Recursive	$\mathcal{O}(n^5 \log n)$	$\mathcal{O}(n^7)$	$\mathcal{O}(n^3)$ {exact size}
Boltzmann	$\mathcal{O}((\log n)^k)$	$\mathcal{O}((\log n)^k)$	$\mathcal{O}(n^2)$ {exact size}
			$\mathcal{O}(n)$ {approx. size}

Changing the ratio edges-vertices

- Let $y > 0$ be a fixed real value.
- For $n \neq 1$, let x_n be such that $\mathbf{E}\left(size(\Gamma G^{\bullet\bullet\bullet}(x_n, y))\right) = n$

Result: There exists a constant $\mu(y) \in (1, 3)$ such that the ratio edges-vertices of the output of $\Gamma G^{\bullet\bullet\bullet}(x_n, y)$ is almost surely equal to $\mu(y)$ when $n \to +\infty$.



Grammar for complexity calculation

Let $\mathcal{C}$ be a class and $x > 0$ a real value

Define $\Lambda C(x)$ as the average number of operations of $\Gamma C(x)$.

- Union:

  $\Gamma C(x) : \text{Bern}\left(\frac{A(x)}{C(x)} \middle| \frac{B(x)}{C(x)}\right) ? : \Gamma A(x) | \Gamma B(x)$

  $$\boxed{\Lambda C(x) = \frac{A(x)}{C(x)} \cdot \Lambda A(x) + \frac{B(x)}{C(x)} \cdot \Lambda B(x)}$$

- Product:

  $\Gamma C(x) : (\Gamma A(x), \Gamma B(x))$.

  $$\boxed{\Lambda C(x) = \Lambda A(x) + \Lambda B(x)}$$

- Set:

  $\Gamma C(x) : \text{Poiss}\left(A(x)\right) \Rightarrow \Gamma A(x)$

  $$\boxed{\Lambda C(x) = E\left(\text{Poiss}\left(A(x)\right)\right) \cdot \Lambda A(x) = A(x) \cdot \Lambda A(x)}$$

Grammar for complexity calculation

Let $\mathcal{C}$ be a class and $x > 0$ a real value

Define $\Lambda C(x)$ as the average number of operations of $\Gamma C(x)$.

Define $\Sigma C(x)$ as the average size of $\Gamma C(x)$.

- Substitution

  $\Gamma C(x)$: replace each atom of $\Gamma A(B(x))$ by $\Gamma B(x)$

  $$\boxed{\Lambda C(x) = \Lambda A(B(x)) + \Sigma A(B(x)) \cdot \Lambda B(x)}$$

- Rejection

  $\mathcal{A} \subset \mathcal{B}$

  $\Gamma A(x)$: do $\gamma \leftarrow \Gamma B(x)$ until $\gamma \in \mathcal{A}$

  $$\boxed{\Lambda A(x) = \frac{B(x)}{A(x)} \cdot \Lambda B(x)}$$

Complexity results

- For all instances of rejection $\mathcal{A} \subset \mathcal{B}$, the acceptance-probability $\frac{A(x_n)}{B(x_n)}$ is bounded away from 0 when n goes to ∞.
- The grammar for calculations implies the following result:
  Theorem: For $n \geq 1$, let x_n be such that the expected size of $\Gamma G^{\bullet\bullet\bullet}(x_n)$ is n. Then:

$$\Lambda G^{\bullet\bullet\bullet}(x_n) = \mathcal{O}(n).$$

Implementation and experimental results

Overview of the implementation

1) Choose a bunch of target-sizes $n = (1000, 10000, 100000, 1000000)$

2) For each target-size n**, compute** x_n **such that** $E(size(\Gamma G^{\bullet\bullet\bullet}(x_n))) = n$ **and evaluate all generating functions of planar graphs at** x_n

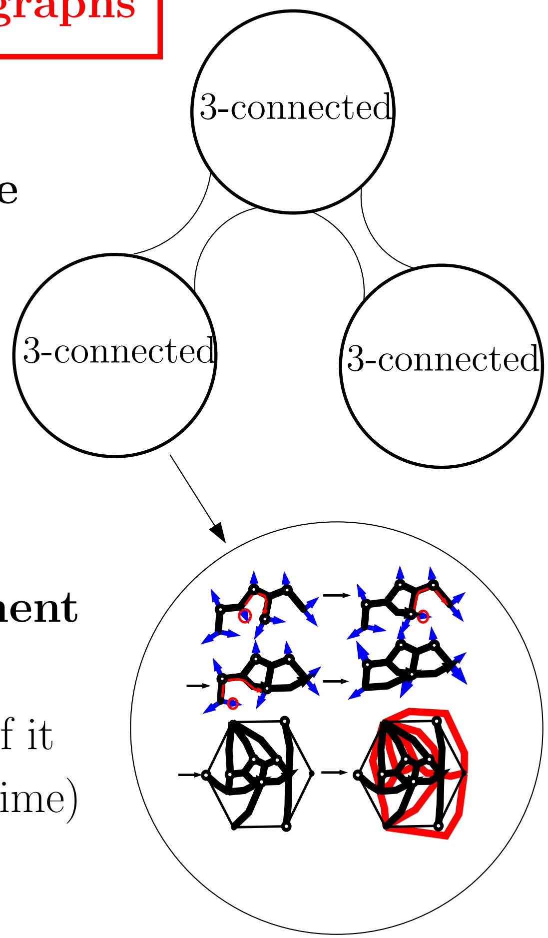
A planar graph is an arborescent structure whose nodes are 3-connected planar graphs

$\Downarrow$

3) Assemble the arborescent structure by doing Bernoulli-choices

$\Rightarrow$ requires ≈ 50 Bernoulli-vectors

Example: choice of the number of connected components with Poiss($C(x_n)$)



4) Generate each 3-connected component

For each node of the arborescent structure generate a binary tree and do the closure of it (the closure can be implemented in linear time)

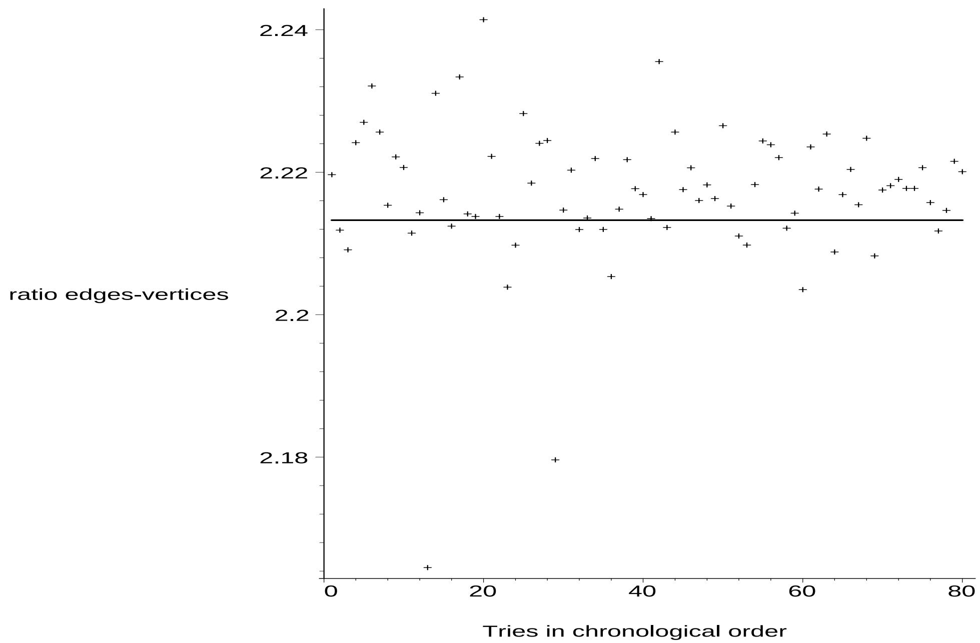
Experimental results

Let X_n be the number of edges of a random planar graph on n vertices.

Theorem: (Giménez, Noy)

There exists a constant $\mu \approx 2.2132$, such that

$$\frac{X_n}{n} \to \mu \quad \text{almost surely when } n \to \infty$$



Experimental results

Conjecture: Let $Y_{n,k}$ be the proportion of vertices having degree k in a random planar graph on n vertices.
Then there is a probability distribution $(p_k)_{k\geq 1}$ such that

$$Y_{n,k} \underset{n\to\infty}{\longrightarrow} p_k \quad \text{almost surely when } n \to \infty$$