Supported by the Austrian Science Fund, Project S8302-MAT.

Thomas Klausner
TU Wien

Joint work with Michel Drmota

TU Wien

Thomas Klausner

INRIA Rocquencourt, December 9, 2002

Patterns in Trees
Outline

• Present assumptions and basic techniques.
• Define patterns.
• Find representations as combinatorial objects.
• Convert to corresponding generating functions.
• Compute asymptotics.
First model: All trees of size $n$ have equal probability.

$$
\frac{iu}{uz} u_d \underbrace{z}_d = (z)d
$$
Explicit number of rooted trees of size $n$ via e.g. Lagrange inversion formula:

$$ (z)d^\mathcal{E} z = \left( \cdots + \frac{i\mathcal{E}}{\mathcal{E}(z)d} + \frac{i\mathcal{Z}}{\mathcal{Z}(z)d} + (z)d + 1 \right) z = (z)d $$

Tree function:

```
...  ...
```

Construction (rooted):  Rooted Trees
Sometimes helpful: planted rooted trees.

Node degree does not change during construction.

Planted Rooted Trees

\[(z)O + z e - 1/2^\theta - 1 = (z)d\]

First order asymptotics

Results in same function \(d(z)\).

\[p(z) = 1 + \frac{p_1}{p_2} z + O(z^{4})\]
Bivariate Generating Functions

Marks special properties with a second variable

E.g.: Nodes with particular degree

\[
\frac{i(1-\gamma)z}{1} + \frac{iz(1-\gamma)z^2}{2(2-\gamma)+1} = \gamma \delta
\]

and \[
\frac{i\gamma^2}{1} = \gamma m \text{ variance, where } \gamma m
\]

Possible to find limit distribution and compute parameters. In this case [Drmota and Gittenberger 1997]: asymptotically Gaussian with mean

\[
\frac{i \gamma (1-n)}{\gamma d} + d \omega z = \frac{i \gamma n}{\gamma d} + \frac{i \gamma}{\gamma d} - d \omega z = d
\]

E.g.: Nodes with particular degree \( k + 1 \)
In our case, connected sub-tree M.

What is a pattern?

Easy example:
How to Mark a Pattern?

More difficult than with single nodes.

Split in parts:

Aim: Describe patterns by generating functions for each part.

\[ \text{System of functional equations.} \]

Split in parts:
Proposition (Planted Rooted Trees)

Let $M$ be a pattern. Then there exist $T$ auxiliary functions $a_j(x;u) (1 \leq j \leq L)$ with $p(x;u) = \sum_{j=1}^{L} a_j(x;u)$ and polynomials $P_j(y_1;...;y_L;u) (1 \leq j \leq L)$ with non-negative coefficients such that

\[ a_1(x;u) = x P_1(a_1(x;u);...;a_L(x;u);u) = (n';x)_P \]

\[ a_L(x;u) = x e^{a_1(x;u)} + \sum_{j=1}^{L-1} x^{L-1-j} P_j(a_1(x;u);...;a_L(x;u);u) = (n';x)_P \]

Let $M$ be a pattern. Then there exist $T$ auxiliary functions (Proposition (Planted Rooted Trees))
Notation to Describe Patterns

- is a (planted) root node, \( \times \) the Cartesian product, \( \cap \) the intersection, \( \cup \) the union, and \( \dot{\cup} \) the disjunct union.

\( \times \) binds stronger than either of \( \cap, \cup, \) and \( \dot{\cup} \).

Note: no immediate one-to-one correspondence to the generating functions (relative probabilities, \( w \)).
\[
\cdot d \times 1 \times \{ \circ \} \quad = \quad 4 \\
\cdot d \times 2 \times \{ \circ \} \quad = \quad 3 \\
\cdot d \times d \times \{ \circ \} \quad = \quad 2 \\
\cdot d \times d \times d \times \{ \circ \} \quad = \quad 1 
\]
3. Convert to Functions

- Standardise no duplicate descriptions of the same structure
- Sprinkle with $n$
- Find coefficients
Each tree \( \mathcal{a} \) represented as disjoint union of trees of the kind

\[
\ldots \times \eta_d^{p} \times \eta_d \times \{\circ\}
\]

(2)

\( (\mathcal{a} \mathfrak{v}) \) (degree of the root of \( \mathcal{a} \)).

Standard Form
Standardising the Functions

Build intersections (symbolically):

\[(\bar{\mathcal{Z}} x \cup \bar{\mathcal{I}} x) \setminus \bar{\mathcal{Z}} x \cap \bar{\mathcal{I}} x = \bar{\mathcal{Z}} x \cap \bar{\mathcal{I}} x\]

\[\{i_{m}^{\ldots i_{1}}\} = \{i_{m}^{\ldots i_{1}}u\} \]

\[\{i_{m}^{\ldots i_{1}}\} = \{i_{m}^{\ldots i_{1}}w\}\]

\[(\bar{\mathcal{Z}} x \cup \bar{\mathcal{I}} x) \times \ldots \times (\bar{\mathcal{Z}} x \cup \bar{\mathcal{I}} x) \times \{0\} \quad \bigcap \quad \bar{\mathcal{Z}} x \times \ldots \times \bar{\mathcal{I}} x \times \{0\} = \bar{\mathcal{Z}} x \cup \bar{\mathcal{I}} x\]

\[\bar{\mathcal{I}} x \times \ldots \times \bar{\mathcal{I}} x \times \{0\} = \bar{\mathcal{I}} x\]

\[\bar{\mathcal{I}} x \times \ldots \times \bar{\mathcal{I}} x \times \{0\} = \bar{\mathcal{I}} x\]

Build intersections (symbolically):
\[ t_2 \times t_1 \times \{\circ\} = \]
\[ = t_2 \times t_1 \times \{\circ\} \cap d \times (t_2 \cup t_1) \times \{\circ\} = 4 \times 3 \]

Only one non-empty intersection:
\[ d \times t_1 \times \{\circ\} = 4 \]
\[ d \times t_2 \times \{\circ\} = 3 \]
\[ d \times d \times d \times \{\circ\} = 2 \]
\[ d \times d \times d \times d \times \{\circ\} = 1 \]

Intersection Example
Example in Standardised Form
Coefficients are computed by simple combinatorics ($k$ is implicitly given by $I$).

Given $l_i$ sub-trees of type $a_i$, $k$ new occurrences of $M$ ($\eta_j l_i$...$\eta_l l_j$): $\mathcal{L} = \text{number of possible configurations of type (2)}$.
Proposed structure of the system of functional equations (1).

\[ (n', (n' x)^1 T_p' \cdots (n' x)^L T_p') P \cdot x = (n' x)^P \]

\[ (I' \tau_i' \cdots \tau_i') P \underbrace{V_{1-L}}_{1-I} - \tau_i + \cdots + \tau_i = (\tau_i' \cdots \tau_i') P \]

\[ I - T \geq l \geq 1 \quad \sum_{I} (n' \tau_i' \cdots \tau_i') P \underbrace{V_{1-L}}_{1-I} = (n' \tau_i' \cdots \tau_i') P \]
How to Find $k = k(l_1; \ldots; l_L)$

New patterns occur when all necessary sub-trees are attached to an node of proper degree.

3. Node of degree five with a $t_3$ attached. Each $t_3$ produces another pattern.

2. Node of degree four with a $t_4$ attached. Each $t_4$ produces another pattern.

1. Node of degree three with a $t_1$ and $t_2$.

In example, three cases:

- Node of proper degree.
- New patterns occur when all necessary sub-trees are attached to

$$(l_1, l_2, \ldots)$$

How to Find $\gamma = \gamma(l_1, l_2, \ldots)$
\[
\frac{i \tau}{\varepsilon(6\hbar^2 + 5\hbar^2 + 4\hbar + 3\hbar n + 2\hbar n + \hbar n)} = \phi_d
\]
\[
\frac{i \tau}{4(6\hbar^2 + 5\hbar^2 + 4\hbar + 3\hbar n + 2\hbar n + \hbar n)} = \varphi_d
\]
\[
\left(\frac{2i}{\hbar}\right)^4 + (6\hbar + 3\hbar + 2\hbar + \hbar) = \gamma_d
\]
\[
\left(\frac{2i}{\hbar}\right)^5 + (6\hbar + 3\hbar + 2\hbar + \hbar) = \delta_d
\]
\[
\begin{align*}
\cdot \left( \sum_{\varphi=1}^{\infty} \int_{9}^{1} i u \ \frac{g-u}{\varphi} \right) x + \\
\frac{i \varphi}{\varepsilon \left( \varepsilon + \varphi + \varphi + 1 \right)} x + \frac{i \varphi}{\varepsilon \left( \varepsilon + \varphi + \varphi + 1 \right)} x + x = 9 \varphi = (n'x)^{9 \varphi}
\end{align*}
\]
Strong Connectivity

Each $a_i$ depends on $a_6$ either directly or through a chain to a leaf (see pattern).

$a_6$ depends on itself and all others (last term).
There exists an analytic function $G(x; u; a_1, \ldots, a_L)$ with non-negative Taylor coefficients such that

$$r(x; u) = G(x; u; a_1(x; u), \ldots, a_L(x; u)).$$

where the $a_i$ were defined earlier.

$$(n' x)^T a \cdots (n' x)^T a n' x G = (n' x) t$$

Proposition (Rooted Trees)
Counting Patterns in Rooted Trees

For the example:

\[
\begin{align*}
  r_0 &= \exp \frac{5x}{5!} \exp \frac{4x}{4!} \exp \frac{3x}{3!} \quad : \quad 0x
\end{align*}
\]

Marks in sub-trees already counted correctly, only have to dis-
tribute marks for newly appearing patterns.

\[
(\text{"uninteresting" sub-trees})
\]

\[
\frac{i}{3} \frac{d}{dx} - \frac{i}{4} \frac{d}{dx} - \frac{i}{5} \frac{d}{dx} - i^{\partial x} = 0x
\]

For the example:

Counting Patterns in Rooted Trees
Newly Appearing Patterns
\text{Unrooted Trees}

\frac{u/m}{u/m} = \frac{m}{m}
4. Asymptotics

Use Drmota's Theorem on systems of functional equations.


Finding the singularity

Often the difficult part:

Paraphrased: Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$. Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$.

Paraphrased: Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$. Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$. Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$. Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$. Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$. Under certain conditions for the system of equations, the coefficients asymptotically follow a Gaussian distribution with mean and variance asymptotically proportional to $n$.
For Our Case

Singularity known:

To find: left eigenvector (for the eigenvalue one) of the derivative of the functional matrix with respect to the functions.

Application of theorem then gives expectation for examples as

\[ \frac{12e(576e^3 + 24e^2 - 25)}{384e - 19} = 0.0026803 \ldots \]

\[ e \]

\[ \begin{array}{c} \end{array} \]

For Our Case
Different kind of trees or tree distributions

Possible future extensions

Dierent kind of trees or tree distributions

Does a 0-1-law hold?

Logical terms using \( \land \), \( \lor \), \( \exists \), and \( \forall \) and describing more general patterns.