Non-contiguous pattern avoidance in binary trees

Michael Dairyko (Pomona College)
Lara Pudwell (Valparaiso University)
Samantha Tyner (Augustana College/Iowa State)
Casey Wynn (Hendrix College/Kent State)

Permutation Patterns 2012
June 15, 2012

Partially supported by NSF grant DMS-0851721

Key Question

How many permutations of length n avoid a given permutation pattern?

Key Question

How many binary trees with n leaves avoid a given tree pattern?

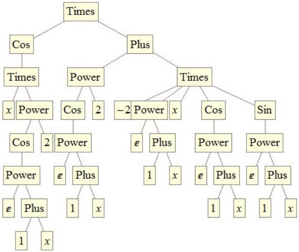
Key Question

How many binary trees with n leaves avoid a given tree pattern?

Concerned with rooted, ordered, full binary trees (each vertex has exactly 0 or 2 children)

History of Tree Patterns: Labelled Trees

- 1983: Flajolet and Steyaert
  - focus on asymptotic probability of avoiding a given pattern
- 1990: Flajolet, Sipala, and Steyaert
  - every leaf of pattern must be matched by a leaf of the tree
  - motivated by compactly storing expressions in computer memory
  - e.g. $\dfrac{d}{dx}\left(\sin(x\cos^2(e^{x+1}))\right) =$



History of Tree Patterns: Labelled Trees

- 1983: Flajolet and Steyaert
  - focus on asymptotic probability of avoiding a given pattern
- 1990: Flajolet, Sipala, and Steyaert
  - every leaf of pattern must be matched by a leaf of the tree
  - motivated by compactly storing expressions in computer memory
- 2012: Dotsenko
  - pattern may occur anywhere in tree
  - motivated by operad theory

History of Tree Patterns: Unlabelled Trees

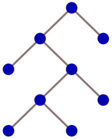
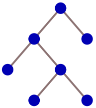
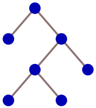
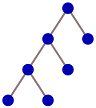
- 2009: Rowland
  - contiguous pattern avoidance in binary trees
  - patterns can be anywhere, not just at leaves
- 2010: Gabriel, Peske, P., Tay
  - extended Rowland's results to m-ary trees
- 2011: Dairyko, P., Tyner, Wynn
  - non-contiguous pattern avoidance in binary trees

Tree patterns

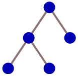
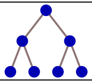
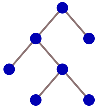
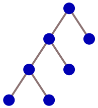
Contiguous tree pattern (Rowland)

Tree T contains tree t if and only if T contains t as a contiguous rooted ordered subtree.

Example:

 contains  and  but avoids .

Contiguous pattern enumeration data

Pattern t	Number of n leaf trees avoiding t
	0
	$\begin{cases} 1 & n=1 \\ 0 & n>1 \end{cases}$
	1
	2^{n-2}
	2^{n-2}
	M_{n-1} (Motzkin numbers)

Contiguous tree pattern enumeration

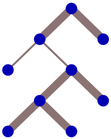
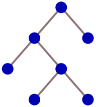
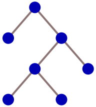
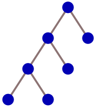
Rowland

- Devised algorithm to find functional equation for avoidance generating function for any set of tree patterns.
- Generating functions are always algebraic.
- Enumerated trees containing specified number of copies of a given tree pattern.
- Completely determined Wilf classes for tree patterns with at most 8 leaves.

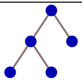
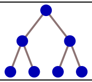
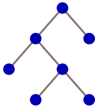
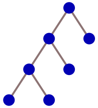
Non-contiguous tree pattern (Dairyko, P., Tyner, Wynn)

Tree T contains tree t if and only if there exists a sequence of edge contractions of T that produce T^* which contains t as a contiguous rooted ordered subtree.

Example:

 contains , , and .

Non-contiguous pattern enumeration data

Pattern t	Number of n leaf trees avoiding t
	0
	$\begin{cases} 1 & n=1 \\ 0 & n>1 \end{cases}$
	1
	2^{n-2}
	2^{n-2}
	2^{n-2}

The Main Theorem

Notation

- Let $\mathrm{av}_t(n)$ be the number of n-leaf trees that avoid t non-contiguously.
- Let $g_t(x) = \sum_{n=1}^{\infty} \mathrm{av}_t(n) x^n$.

The Main Theorem

Notation

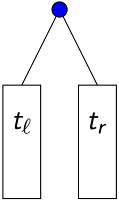
- Let $\text{av}_t(n)$ be the number of n-leaf trees that avoid t non-contiguously.
- Let $g_t(x) = \sum_{n=1}^{\infty} \text{av}_t(n) x^n$.

Theorem

Fix $k \in \mathbb{Z}^+$. Let t and s be two k-leaf binary tree patterns. Then $g_t(x) = g_s(x)$.

(More) Notation

- Given tree t,
  - let t_ℓ be the subtree whose root is the left child of t's root.
  - let t_r be the subtree whose root is the right child of t's root.



Notation and Computation

(More) Notation

- Given tree t,
  - let t_ℓ be the subtree whose root is the left child of t's root.
  - let t_r be the subtree whose root is the right child of t's root.

Notice

$$g_t(x) = x + g_{t_\ell}(x) \cdot g_t(x) + g_t(x) \cdot g_{t_r}(x) - g_{t_\ell}(x) \cdot g_{t_r}(x)$$

Notation and Computation

(More) Notation

- Given tree t,
  - let t_ℓ be the subtree whose root is the left child of t's root.
  - let t_r be the subtree whose root is the right child of t's root.

Notice

$$g_t(x) = x + g_{t_\ell}(x) \cdot g_t(x) + g_t(x) \cdot g_{t_r}(x) - g_{t_\ell}(x) \cdot g_{t_r}(x)$$

Solving...

$$g_t(x) = \frac{x - g_{t_\ell}(x) \cdot g_{t_r}(x)}{1 - g_{t_\ell}(x) - g_{t_r}(x)}.$$

Proposition

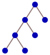
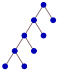
$$g_t(x) = \frac{x - g_{t_\ell}(x) \cdot g_{t_r}(x)}{1 - g_{t_\ell}(x) - g_{t_r}(x)}.$$

Proposition

For any tree pattern t, $g_t(x)$ is a rational function of x.

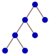
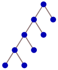
A special case...

Let c_k be the k-leaf left comb
(the unique k-leaf binary tree where every right child is a leaf).

$c_1 =$ , $c_2 =$ , $c_3 =$ , $c_4 =$ , $c_5 =$ ,etc.

A special case...

Let c_k be the k-leaf left comb
(the unique k-leaf binary tree where every right child is a leaf).

$c_1 = \bullet$, $c_2 =$ , $c_3 =$ , $c_4 =$ , $c_5 =$ ,etc.

If $t = c_k$, then $t_\ell = c_{k-1}$ and $t_r = \bullet$.

For $k \geq 2$, we have

$$g_{c_k}(x) = \frac{x - g_{c_{k-1}}(x) \cdot g_{\bullet}(x)}{1 - g_{c_{k-1}}(x) - g_{\bullet}(x)} = \frac{x}{1 - g_{c_{k-1}}(x)}.$$

Back to the main result

Theorem

Fix $k \in \mathbb{Z}^+$. Let t and s be two k-leaf binary tree patterns. Then $g_t(x) = g_s(x)$.

Proof sketch

Inductive step:

- Assume the theorem holds for tree patterns with ℓ leaves where $\ell < k$.
- Then any ℓ-leaf tree has avoidance generating function $g_{c_\ell}(x)$.
- Consider tree t with ℓ leaves to the left of its root and tree s with $\ell + 1$ leaves to the left of its root.
- Do algebra with previous work to show that $gf_t(x) = gf_s(x)$.

Generating functions

k	$g_{c_k}(x)$	OEIS number
1	0	trivial
2	x	trivial
3	$\frac{x}{1-x}$	trivial
4	$\frac{x-x^2}{1-2x}$	A000079
5	$\frac{x-2x^2}{1-3x+x^2}$	A001519
6	$\frac{x-3x^2+x^3}{1-4x+3x^2}$	A007051
7	$\frac{x-4x^2+3x^3}{1-5x+6x^2-x^3}$	A080937
8	$\frac{x-5x^2+6x^3-x^4}{1-6x+10x^2-4x^3}$	A024175
9	$\frac{x-6x^2+10x^3-4x^4}{1-7x+15x^2-10x^3+x^4}$	A080938

An explicit formula

Theorem

Let $k \in \mathbb{Z}^+$ and let t be a binary tree pattern with k leaves. Then

$$g_t(x) = \frac{\sum_{i=0}^{\lfloor \frac{k-2}{2} \rfloor} (-1)^i \cdot \binom{k-(i+2)}{i} \cdot x^{i+1}}{\sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} (-1)^i \cdot \binom{k-(i+1)}{i} \cdot x^i}.$$

...and permutations

We know that the Catalan numbers count:

- the number of binary trees
- the number of 231-avoiding permutations

Can we say more?

...and permutations

We know that the Catalan numbers count:

- the number of binary trees
- the number of 231-avoiding permutations

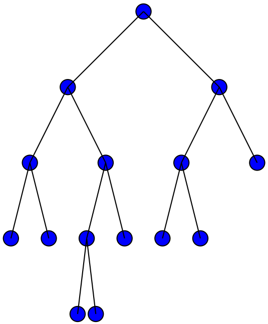
Can we say more?

Theorem

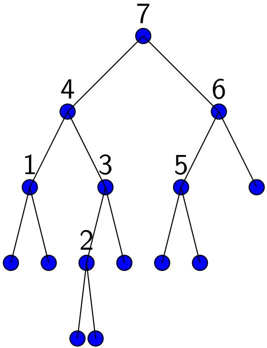
Let t be any binary tree pattern with $k \geq 2$ leaves. Then

$$\mathrm{av}_t(n) = \mathrm{s}_{n-1}(231, (k-1)(k-2)\cdots 21).$$

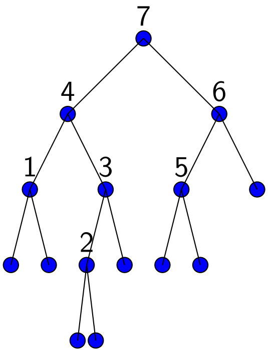
Example



Example



Example



1423756

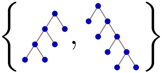
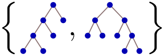
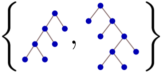
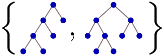
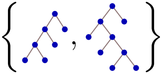
Avoiding multiple tree patterns

- Methods extend naturally to trees avoiding multiple tree patterns simultaneously:
  - Generating functions are still rational.

Avoiding multiple tree patterns

- Methods extend naturally to trees avoiding multiple tree patterns simultaneously:
  - Generating functions are still rational.
  - No longer one Wilf class per size of tree pattern

Wilf classes for avoiding a 4 leaf and a 5 leaf tree pattern

Pattern representatives	OEIS
	0 for $n \geq 11$
	A016777 $(3k+1)$
	A152947 $\left(\frac{(k-2)\cdot(k-1)+1}{2}\right)$
	A000071 (Fibonacci numbers -1)
	A000073 (Tribonacci Numbers)

Avoiding multiple tree patterns

- Methods extend naturally to trees avoiding multiple tree patterns simultaneously:
  - Generating functions are still rational.
  - No longer one Wilf class per size of tree pattern (Open: Find a combinatorial characterization of when two sets of tree patterns are Wilf equivalent.)

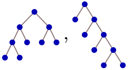
Avoiding multiple tree patterns

- Methods extend naturally to trees avoiding multiple tree patterns simultaneously:
  - Generating functions are still rational.
  - No longer one Wilf class per size of tree pattern (Open: Find a combinatorial characterization of when two sets of tree patterns are Wilf equivalent.)
  - Some sets of patterns have enumeration sequences that obviously count a set of pattern-avoiding permutations. Others clearly aren't (classical) permutation sequences.

Avoiding multiple tree patterns

- Methods extend naturally to trees avoiding multiple tree patterns simultaneously:
  - Generating functions are still rational.
  - No longer one Wilf class per size of tree pattern (Open: Find a combinatorial characterization of when two sets of tree patterns are Wilf equivalent.)
  - Some sets of patterns have enumeration sequences that obviously count a set of pattern-avoiding permutations. Others clearly aren't (classical) permutation sequences. Example:

$$\left\{ \mathrm{av}_{\{ T_1, T_2 \}}(n) \right\}_{n=2}^{\infty} = 1, 2, 5, 12, 26, 49, 83, 129, \ldots$$



Valparaiso University

Avoiding multiple tree patterns

- Methods extend naturally to trees avoiding multiple tree patterns simultaneously:
  - Generating functions are still rational.
  - No longer one Wilf class per size of tree pattern (Open: Find a combinatorial characterization of when two sets of tree patterns are Wilf equivalent.)
  - Some sets of patterns have enumeration sequences that obviously count a set of pattern-avoiding permutations. Others clearly aren't (classical) permutation sequences. (Open: Precisely characterize which sets of tree patterns correspond to classical permutation sequences.)

Avoiding multiple tree patterns

- Methods extend naturally to trees avoiding multiple tree patterns simultaneously:
  - Generating functions are still rational.
  - No longer one Wilf class per size of tree pattern (Open: Find a combinatorial characterization of when two sets of tree patterns are Wilf equivalent.)
  - Some sets of patterns have enumeration sequences that obviously count a set of pattern-avoiding permutations. Others clearly aren't (classical) permutation sequences. (Open: Precisely characterize which sets of tree patterns correspond to classical permutation sequences.)

    (Open: Let f be the vertex-labelling bijection between binary trees and 231-avoiding permutations given before. Let S be a set of tree patterns. Characterize which permutations correspond to S-avoiding trees under f.)

Summary

- $g_t(x)$ is rational and of a very nice form for any non-contiguous tree pattern t.
- Only one Wilf class for each number of leaves!
- Trees avoiding a k-leaf tree pattern are in bijection with permutations avoiding 231 and $(k-1)(k-2)\cdots 1$.
- Several open questions remain for trees avoiding *sets* of non-contiguous patterns.

Thank You!

References

- M. Dairyko, L. Pudwell, S. Tyner, and C. Wynn, Non-contiguous pattern avoidance in binary trees, preprint, `http://arxiv.org/abs/1203.0795`
- V. Dotsenko, Pattern avoidance in labelled trees, *S'em. Lothar. Combin.*, B67b (2012), 27 pp.
- P. Flajolet, P. Sipala, and J. M. Steyaert, Analytic variations on the common subexpression problem, *Automata, Languages, and Programming: Proc. of ICALP 1990*, Lecture Notes in Computer Science, Vol. 443, Springer, 1990, pp. 220–234.
- N. Gabriel, K. Peske, L. Pudwell, and S. Tay, Pattern avoidance in ternary trees, *J. Integer Seq.* **15** (2012), 12.1.5.
- E. S. Rowland, Pattern avoidance in binary trees, *J. Combin. Theory, Ser. A* **117** (2010), 741–758.
- J. M. Steyaert and P. Flajolet, Patterns and pattern-matching in trees: an analysis, *Info. Control* **58** (1983), 19–58.