Beurling Integers: Part 4

Analysis of the Beurling Tree

Devin Platt

July 11, 2015

In this post we analyze the tree representing the partial order of Beurling integers over factorizations. We do this by building a number triangle containing the frequencies of occurrence of prime counting functions.

A major object in the study of the primes is the prime counting function. Beurling integers can really have any prime counting function since what numbers are prime is completely arbitrary. For any given height h of the tree, there are $h+1$ choices for the value of the prime counting function: $1, 2, 3, ...h+1$. However, the number of leaves grows exponentially with h, so a natural question to ask is how this exponential number of paths is distributed among a linear number of options for the value of the prime counting function. This question can be encapsulated in a function. Let

$$T(h, p-1) = \text{the number of paths down the tree of height h with p primes.}$$

We can build a number triangle from this function much like how Pascals triangle is built from binomial coefficients:

T(0,0)				
T(1,0)	T(1,1)			
T(2,0)	T(2,1)	T(2,2)		
T(3,0)	T(3,1)	T(3,2)	T(3,3)	
T(4,0)	T(4,1)	T(4,2)	T(4,3)	T(4,4)

The values of this triangle are:

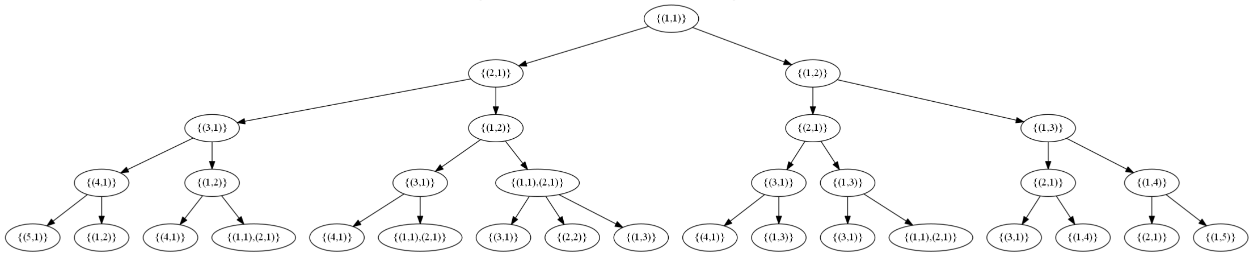
1				
1	1			
1	2	1		
1	3	3	1	
1	5	6	4	1

These values can be verified with the tree.

Some properties of the triangle:

- $T(h, p-1) = 0$ when $h < 0$, $p-1 < 0$, or $p-1 > h$.
- $T(h, 0) = 1$
- $T(h+1, p-1) \geq T(h, p-1) + T(h, p-2)$ for $p \geq 2$.
- $T(h, p-1) < T(h+1, p-1)$ when $p-1 > 0$. (Columns are increasing)

Figure 1: The Beurling Tree



- $T(h, p-1) \geq C(h, p-1)$, where $C(n, k)$ is the binomial coefficient.

The first property is simply stating the bounds of T. The second property comes from the fact that there is only one sequence with only one prime: the exponential sequence. The third property comes from the fact that each node has one prime child and at least one composite child. The fourth and fifth properties follow from the third. For the fifth property recall that Pascal's triangle has the relationship $C(n, k) = C(n-1, k) + C(n-1, k-1)$. Essentially, if the tree were binary this triangle would be Pascals. Note that unlike Pascals triangle, this triangle is NOT symmetric.

Lets observe a larger triangle:

1												
1	1											
1	2	1										
1	3	3	1									
1	5	6	4	1								
1	6	14	10	5	1							
1	7	22	30	15	6	1						
1	9	31	59	55	21	7	1					
1	11	45	103	130	91	28	8	1				
1	12	71	172	276	251	140	36	9	1			
1	15	90	319	537	632	441	204	45	10	1		
1	16	120	491	1134	1425	1288	722	285	55	11	1	
1	17	156	731	2074	3390	3325	2402	1119	385	66	12	1

It appears that the nth diagonal is an n-1 degree polynomial:

Degree	Sequence	Polynomial
0	1,...	1
1	1,2,...	x
2	1, 3, 6 ...	$x/2 + x^2/2$
3	1, 5, 14, 30, ...	$x/6 + x^2/2 + x^3/3$
4	1, 6, 22, 59, 130, ...	$7x/12 - x^2/8 + 5x^3/12 + x^4/8$
5	1, 7, 31, 103, 276, 632, ...	$-(13x)/60 + (31x^2)/24 - x^3/3 + (5x^4)/24 + x^5/20$
6	1, 9, 45, 172, 537, 1425, 3325, ...	$(2x)/5 - (41x^2)/60 + (3x^3)/2 - x^4/3 + x^5/10 + x^6/60$

We can observe this by taking the first n numbers of the nth diagonal and interpolating them as a polynomial. We can then verify this formula for the diagonal with later numbers in the diagonal, though this is not a proof of the formulas validity.

Such is also the case with Pascals triangle, where polynomial formulas for the diagonals are given using the Stirling numbers of the First Kind to determine the coefficients. If we prove that this triangle has diagonal formulas and find a general formula, we would have a formula for $T(h, p-1)$.

The first three diagonals are the same as those in Pascals triangle: 1s, x, and and the triangular numbers. The fourth is different, it is the square pyramidal numbers. Unfortunately, a search for the numbers of the fifth diagonal (or later ones) in the Online Encyclopedia of Integer Sequences returns nothing. I have not been able to identify any patterns in the coefficients of the polynomials either.

So we move on to how to prove a formula for a single diagonal. This turns out to be a systematic procedure.

First note that the nth diagonal gives $T(h, h-n+1)$. It gives the number of paths which have $n-1$ composites for a given height. Note that the nth diagonal begins on row n. The first diagonal is 1's because there is only one way to have a sequence of all primes.

For the second diagonal recognize that the first composite number must be the first prime squared. Thus we only need to count the options where this number could be. The second diagonal is x because there are x ways of choosing one composite out of x numbers. For the third diagonal, consider the two possibilities for our second composite number: (1,3) and (1,1),(2,1). For the first possibility we have:

$$\{(1,1)\}, \{(1,2)\}, \text{ some primes}, \{(1,3)\}, \text{ more primes}, \ldots$$

Because this is the third diagonal, we disregard the first two numbers (this is simply how our polynomial is offset because it starts on the third row.) Thus we see that we have $C(x,1) = x$ choices.

For the second possibility we have:

$$\{(1,1)\}, \{(2,1)\}, \text{ some primes}, \{(1,2)\}, \text{ some primes}, \{(1,1),(2,1)\}$$

So we have $C(x,2) = x(x-1)/2$ choices.

So in total we have that the diagonal is $C(x,1) + C(x,2) = x(x+1)/2$.

The general algorithm is as follows:

- Build a modified version of the tree in which each path to a leaf has n-1 primes and n-1 composites.
- Traverse this tree to find all of the possible sequences of $(n-1)$ composites.
- Each of these sequences of composites will have a binomial coefficient associated with it which will be part of the sum that makes the diagonal formula.
- There is one special case: a single branch with $n-1$ primes followed by $n-1$ composites, which are each the first prime times a prime number. Its associated coefficient is $C(x, n-1)$.
- For each of these sequences of composites (besides the special case), find the lowest node that is in each branch with this sequence of composites. Call this node n and its height of this node h.
- Note the number of elements of the sequence of composites which are below n, and call this number of remaining composites r.

- The associated binomial coefficient with this sequence of composites is $C(x-(h+1)+(n-1), r)$. That is to say that we choose the placement of r composites into x slots, minus the $h+1$ slots that are taken up by a predetermined ordering, plus $n-1$ slots because our diagonal starts on row $n-1$. This only works if $h+1 < n-1$, hence why our special case is needed to be taken care of separately.

Proof of correctness of algorithm

For the most part, the algorithm really is its own proof. It looks at all distinct possibilities for having a subsequence of $n-1$ composites, then counts the number of ways of achieving each. The main method works when $h+1 < n-1$, so what needs to be shown is that there is only one exception. After showing this, we prove that the exception always exists and has a coefficient of $C(x, n-1)$.

The composite subsequence which is: (1,2), (1,1),(2,1), (1,1),(3,1), has only the full sequence: primes, (1,2), (1,1),(2,1), (1,1),(3,1), since the introduction of a composite before the (n-1)th prime would change the sequence of composites. Thus for this sequence we would obtain $h+1 = 2(n-1)$, which invalidates the main method.

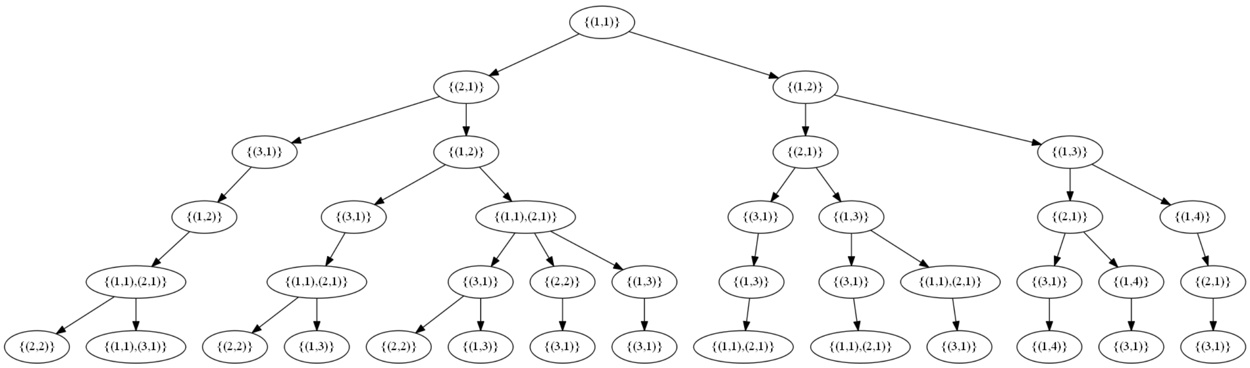
Now we consider two cases: other sequences of composites following $n-1$ primes, and any other subsequence of composites which is contained within a sequence that has at least one composite before the (n-1)th prime.

Now any other subsequence of composites following the (n-1) primes contains $(2,2)$, since $(2,2)$ is must be the first composite to break this chain of the first prime times other primes. More importantly, it does not contain (1,1),(n-1,1) because that would require (1,1),(n-2,1) to come first and so on and would imply our original special case of a sequence containing only the first prime times primes. The branch of the tree that has: (n-2) primes, (1,2), (n-1,1), will have a path with our subsequence of composites because all that weve done is switch the positions of (1,2) and (n-1,1), which only invalidates the possibility of having (1,1),(n-1,1) in our subsequence of composites (due to translation invariance).

Now suppose that our subsequence of composites is contained in a path which has at least one composite occur before the (n-1)th prime. Consider the last prime which occurs immediately following a composite. There exists a path in the tree which is exactly the same except that these two numbers are swapped in order. Its easy to see that this swapping wont affect anything with translation invariance since the composite is at a position greater than or equal to n-1 in the sequence.

An example of the algorithm for the fourth diagonal

The tree generated:



The coefficients corresponding to each subsequence of composite numbers:

{(1,2)},{(1,1),(2,1)},{(2,2)}	$C(x+1,3)$
{(1,2)},{(1,1),(2,1)},{(1,1),(3,1)}	$C(x,3)$
{(1,2)},{(1,1),(2,1)},{(1,3)}	$C(x,2)$
{(1,2)},{(1,3)},{(1,1),(2,1)}	$C(x,2)$
{(1,2)},{(1,3)},{(1,4)}	$C(x,1)$

You may verify that these values agree with the algorithm as described, and that the sum of these binomial coefficients yields the correct polynomial. This algorithm is implemented on my Github project for the Beurling integers (https://github.com/devinplatt/BeurlingTree). Using the program Ive calculated the formulas in the sum of binomial coefficient form. A list of the first four coefficient formulas (starts with second diagonal):

C(x, 1)
C(x, 1)+ C(x, 2)
C(x, 1)+ C(x, 2)+ C(x, 2)+ C(x, 3)+ C(x+1, 3)
C(x, 1)+ C(x, 2)+ C(x, 2)+ C(x, 3)+ C(x, 3)+ C(x, 3)+ C(x, 4)+ C(x+1, 3)+ C(x+1, 3)+ C(x+1, 4)+ C(x+1, 4)

I have not yet been able to find a pattern in these formulas, except that each diagonal sum contains the summands of the last, which might imply some sort of recursive relation.

Note that the existence of our algorithm implies that the diagonals are polynomial. The fact that the nth diagonal contains a factor $C(x, n-1)$ and no factor with a greater value as its second argument implies that the nth diagonal is an $n-1$ degree polynomial. This means that simply using polynomial interpolation is a perfectly valid technique for determining the formulas (though it may be easier to prove a general formula with the sum of binomial coefficients form). I wish that I could provide a general formula for this triangle, but at the moment I do not have one. This article concludes my series on the Beurling integers. I will continue to write about them, but building the tree and analyzing the triangle has been the bulk of my work, and my purpose in this series was to introduce them.