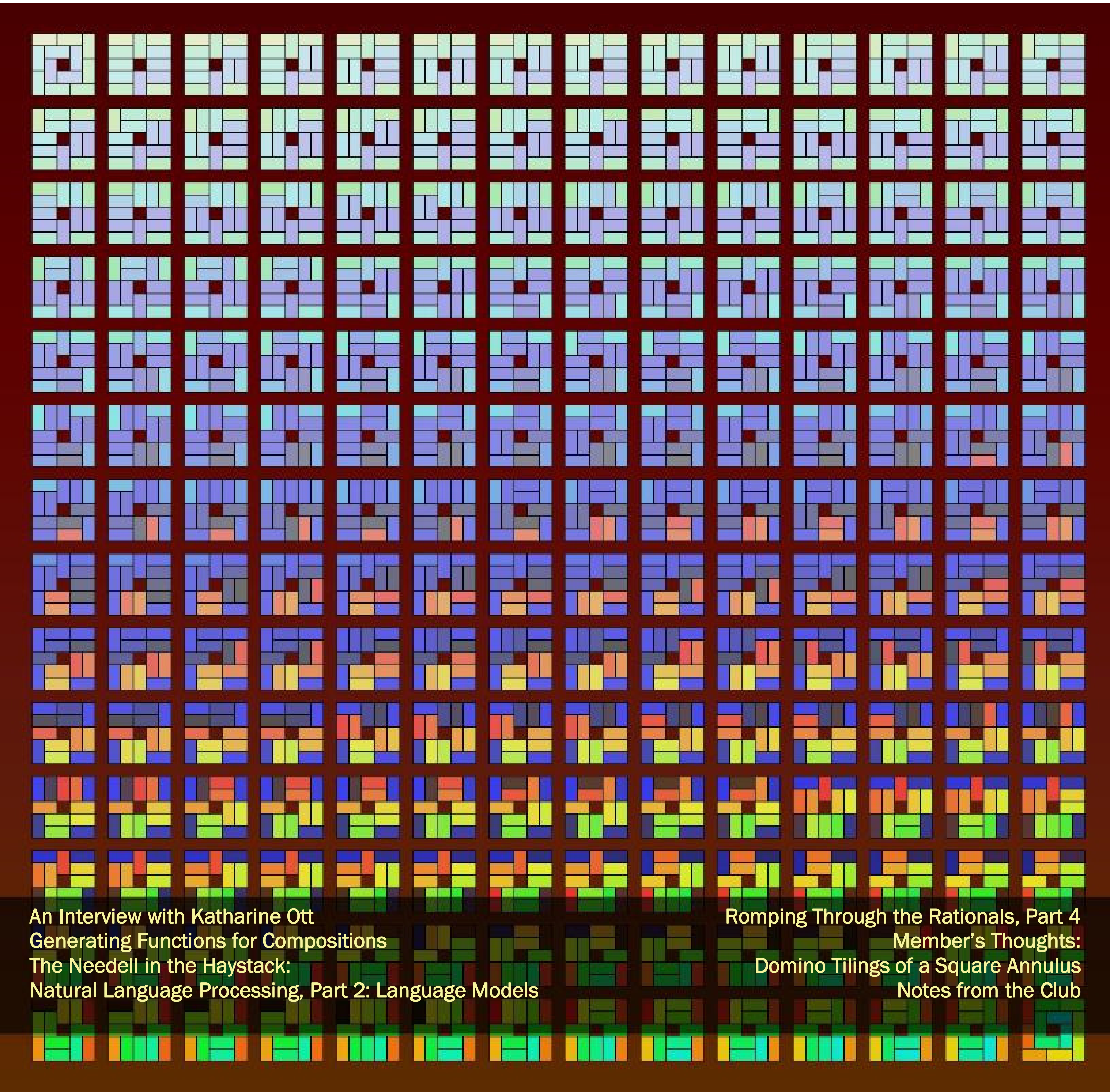
Girls' Angle Bulletin

April/May 2023 • Volume 16 • Number 4

To Foster and Nurture Girls' Interest in Mathematics



From the Founder

If you're interested in doing math research, get into the habit of playing with math. As you play with math, more questions will occur to you, and these questions will induce further play. Before you know it, you'll be playing with something new. - Ken Fan, President and Founder

Girls' Angle Donors

A Heartfelt Thank You to our Donors!

Individuals

Uma Achutha
Nancy Blachman and David desJardins, founders of the Julia Robinson Mathematics Festival, jrmf.org.
Michael Beckman
Ravi Boppana
Marla Capozzi
Yungyuan Chang
Suhas Daftuar
Patricia Davidson
Ingrid Daubechies
Anda Degeratu
Ellen Eischen
Glenn and Sara Ellison
John Engstrom
Lena Gan
Jacqueline Garrahan
Courtney Gibbons
Shayne Gilbert
Vanessa Gould
Rishi Gupta
Larry Guth
Andrea Hawksley
Scott Hilton
Delia Cheung Hom and Eugene Shih
David Kelly
Mark and Lisel Macenka
Brian and Darlene Matthews
Toshia McCabe
Mary O'Keefe
Stephen Knight and Elizabeth Quattrocki Knight
Sarah Manchester
Alison and Catherine Miller
Beth O'Sullivan
Robert Penny and Elizabeth Tyler
Malcolm Quinn
Veena Reddy
Jeffrey and Eve Rittenberg
Marcine Snyder
Eugene Sorets
The Waldman and Romanelli Family
Marion Walter
Andrew Watson and Ritu Thamman
Brandy Wiegers
Brian Wilson and Annette Sassi
Lissa Winstanley
Cunyuan Zhang
The Zimmerman family
Anonymous

Nonprofit Organizations

Draper Laboratories
The Mathenaeum Foundation
Orlando Math Circle

Corporate Donors

Adobe
Akamai Technologies
Big George Ventures
D. E. Shaw
John Hancock
Maplesoft
Massachusetts Innovation & Technology Exchange (MITX)
Mathenaeum
MathWorks, Inc.
Microsoft
Microsoft Research
Nature America, Inc.
Oracle

For Bulletin Sponsors, please visit girlsangle.org.

Girls' Angle Bulletin

The official magazine of
Girls' Angle: A Math Club for girls
Electronic Version (ISSN 2151-5743)

Website: www.girlsangle.org
Email: girlsangle@gmail.com

This magazine is published six times a year by Girls' Angle to communicate with its members and to share ideas and information about mathematics.

Girls' Angle welcomes submissions that pertain to mathematics.

The print version of the Bulletin is printed by the American Mathematical Society.

Editors: Amanda Galtman
Jennifer Sidney
Executive Editor: C. Kenneth Fan

Girls' Angle: A Math Club for Girls

The mission of Girls' Angle is to foster and nurture girls' interest in mathematics and empower them to tackle any field no matter the level of mathematical sophistication.

On the cover: All 196 domino tilings of the 5-annulus as defined in *Member's Thoughts* on page 24.

An Interview with Katharine Ott

Most importantly, I think one must have a growth mindset to succeed in math.

Katharine Ott is Associate Professor of Mathematics at Bates College. She went to Middlebury College and earned her PhD in mathematics at the University of Virginia under the supervision of Irina Mitrea.

This interview was conducted by Raegan Phillips and Ken Fan.

Girls' Angle: At what age did you discover your interest in mathematics? What sparked your interest? Was there someone who encouraged your curiosity in mathematics?

Katharine: The first time I have a clear memory of being interested in mathematics is high school. I remember that I always did my math homework first, and sort of looked forward to it (though I wouldn't have said that out loud at the time to anyone!). When I graduated with my PhD, my brother gave me a newspaper article (with photo) from a math competition that I participated in from the 3rd or 4th grade. It's weird that I don't remember those competitions or succeeding in them, but I guess it means that my interest in math started even earlier than high school.

Girls' Angle: Could you please retrace your journey to becoming a professional mathematician? What were the most difficult challenges that you had to overcome and how did you overcome them?

Katharine: Even though I really enjoyed math in high school, I didn't say out loud that I was interested in math until college. I think part of my hesitancy stemmed from just generally being a teenager and not wanting to seem too eager about anything in school. But I also think another part of me wasn't ready to be labeled a "math person". I had a lot of interests (and still do!), but for one reason or another it feels like a math interest trumps all. I took Calculus II and then Linear Algebra in my first year at Middlebury College. It was after the Linear Algebra course that I got really hooked on math. I decided that I wanted to be a math major. For the rest of my time in college I was very inspired and supported by my professors, and I held them in high esteem. I felt that they had the best job in the world, so I decided to apply to graduate school. I made this career decision a little late (during my senior year), so unfortunately I missed out on trying some opportunities such as an REU program. In hindsight, I think that I would have enjoyed and benefited a lot from a research experience and more exposure to math beyond Middlebury. Nonetheless, I did get into a graduate program - ONE program. I accepted and went off to the University of Virginia. Graduate school was really challenging for me. It was a big step up academically. As I mentioned, I went to a liberal arts college and while I got a great education, I did not have as many upper level math classes as some students have entering graduate school. I succeeded by finding helpers, including supportive faculty and especially my PhD advisor, and a close-knit group of friends in the math program. I still consider these friends as some of my best friends. They were always there to provide support and we had a ton of fun outside of work time!

It was also really important that I adjusted my mindset. To be more specific, once I spent a month or so in graduate school I realized just how different a playing field it was even from my so-called elite undergrad college. I had to adjust my expectations from having to be among the

Dear Reader,

We're committed to producing quality math educational content and make every effort to provide this content to you for free.

We are also committed to surviving as a nonprofit!

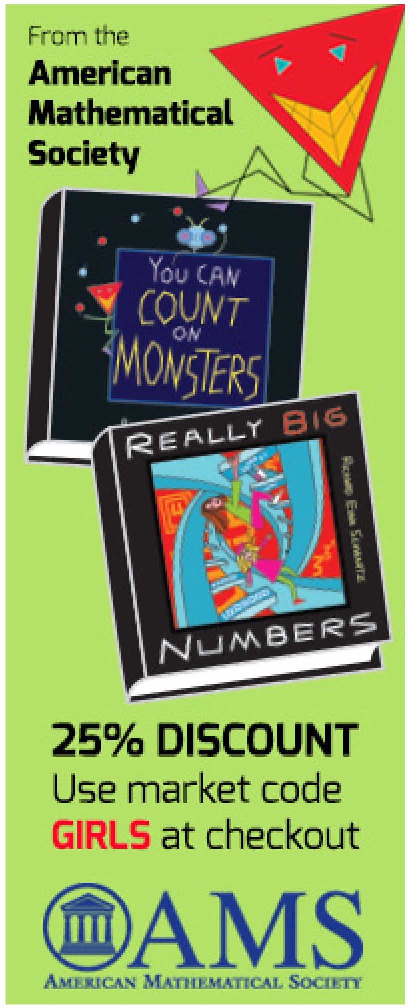
For this issue, those who do not subscribe to the print version will be missing out on a portion of this interview with Prof. Katharine Ott and other content. We hope that you consider the value of such content and decide that the efforts required to produce such content are worthy of your financial support.

We know that mathematical interest and talent is unrelated to economic status, which is why we provide so much content for free. But we hope that those of you who are comfortable financially will help us to continue in our efforts.

So, please consider subscribing to the Bulletin. Thanks to our sponsors, subscriptions cost $36/year. With a subscription, you have also gained access to our mentors via email and the ability to influence content in this Bulletin. Visit www.girlsangle.org/page/bulletin_sponsor.html for more information.

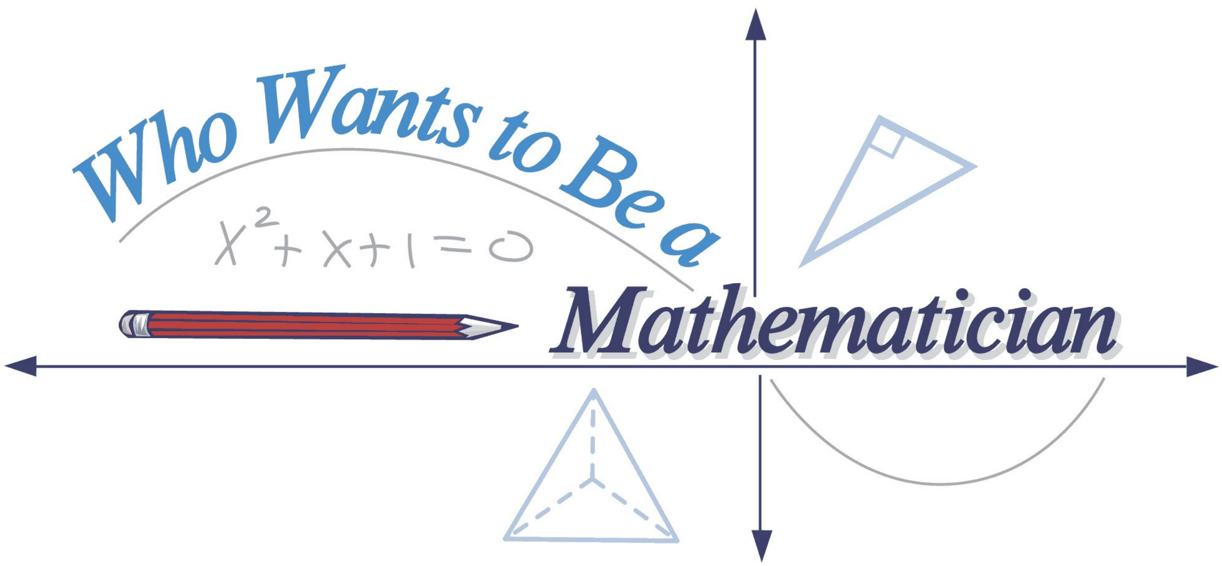
Thank you and best wishes,
Ken Fan
President and Founder
Girls' Angle: A Math Club for Girls

Content Removed from Electronic Version



The American Mathematical Society is generously offering a 25% discount on the two book set *Really Big Numbers* and *You Can Count On Monsters* to readers of this Bulletin. To redeem, go to http://www.ams.org/bookstore-getitem/item=MBK-84-90 and use the code "GIRLS" at checkout.

Content Removed from Electronic Version



America’s Greatest Math Game: Who Wants to Be a Mathematician.

Girls!

Learn Mathematics!

Make new Friends!

Meet Professional Women who use math in their work!

Improve how you Think and Dream!

Girls' Angle

A math club for ALL girls, grades 5-12.

girlsangle@gmail.com
girlsangle.org

Girls' Angle

Content Removed from Electronic Version

Generating Functions for Compositions

by Robert Donley[1]
edited by Amanda Galtman

For this installment, it will be helpful to review the following topics:

Topic	Volume	Issue(s)
Strict and weak compositions, the trinomial formula	16	3
Pascal's triangle, binomial coefficients, hockey stick rule	15	4, 6
Fibonacci numbers and the Binet formula	15	6
Generating functions, geometric power series, binomial series	15	6
Catalan numbers as lattice path or chain counts	15	5

Weak compositions appear in many elementary applications: the binomial theorem, Boolean posets, divisors, and the Fundamental Theorem of Arithmetic. We consider generating functions for compositions to further extend and develop these concepts. With this approach, many major themes and techniques of this series unite, with a central role played by the Fibonacci numbers.

Definition. A **weak composition** of $k \geq 0$ into n **parts** is an ordered sum $x_1 + \ldots + x_n = k$ such that each $x_i \geq 0$. Similarly, for a (strict) **composition** of $k \geq n$ into n **parts**, we require each $x_i > 0$. When convenient, we denote either type by the ordered n-tuple $(x_1, \ldots, x_n)$.

Example: We consider compositions where the parts are either 1 or 2 only, but the number of parts can vary. For instance, such compositions of $k = 3$ are given by

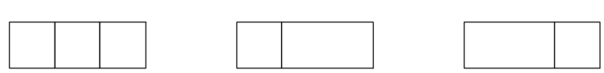
$$1 + 1 + 1 \qquad 1 + 2 \qquad 2 + 1$$

Exercise: Up to $k = 6$, verify that N_k, the number of such compositions (with parts either 1 or 2 only), equals the Fibonacci number F_{k+1}. Recall that the Fibonacci numbers are defined recursively by $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$, for $n \geq 3$. The sequence begins 1, 1, 2, 3, 5, 8, 13, ….

To build on the preceding exercise and prove this equality for larger values of k, we need only show that N_k satisfies the same recursion as the Fibonacci numbers. For $k \geq 3$, every sum has at least two terms. If the sum ends with a 1, there are N_{k-1} ways to begin the sum, and likewise there are N_{k-2} ways if the sum ends with a 2. Since these cases do not overlap and include all sums of interest, we have $N_k = N_{k-1} + N_{k-2}$.

Exercise: How do the equations change if the first term of each composition is 1? If 2?

To visualize this composition model, consider a strip of length k onto which we place squares with side length one or dominos composed of two squares. With $k = 3$, we have the following strips:



[1] This content is supported in part by a grant from MathWorks.

Exercise: Adapt this model to a strip of dimensions 2-by-k onto which we place only dominos. Find the recurrence relation, the first four terms, and the general formula for the number of domino arrangements on a 2-by-k strip. Verify the first four terms by listing all strips.

What if we also allow squares of side length 1? Repeat the previous part and derive the general formula using the method for the Binet formula. (For the answer, look at sequence A015518 in the Online Encyclopedia of Integer Sequences.)

Exercise: With no restrictions on parts, calculate the number of compositions of k by interpreting a strip as a binary word. Specifically, given the word 1010...101 with k 1s, how does removing any number of 0s form a composition of k? Then, with $1 \le n \le k$, verify the formula for the number of compositions of k with n parts from the previous installment.

Let's consider how to encode the above information with a generating function. A composition is an ordered list of parts, so for each position in the list, we assign a variable and form a polynomial by raising that position's variable to successive powers. Now, multiply these polynomials together. Notice how the parts organize properly using exponents; for each term whose exponents sum to k, the exponents correspond to a composition. For instance, to find the compositions of 3 into two parts, we multiply the two polynomials

$$(x_1^1 + x_1^2)(x_2^1 + x_2^2)$$

and keep only the terms whose exponents sum to 3: $x_1^1 x_2^2$ and $x_1^2 x_2^1$. These terms correspond to compositions (1, 2) and (2, 1), respectively. For three parts, we multiply the three polynomials

$$(x_1^1 + x_1^2)(x_2^1 + x_2^2)(x_3^1 + x_3^2).$$

Exercise: Expand the previous expression and find the number of all compositions of 4 into three parts, with those parts restricted to values 1 and 2. What products give the other two compositions when $k = 4$?

To generalize this method, we have an

Algorithm for generating compositions of k:

1. If the number of parts is n, then use the variables $x_1, \ldots, x_n$.
2. If the parts are restricted to $a, b, c, \ldots$, then use factors $x_i^a + x_i^b + x_i^c + \ldots$.
3. Multiply all n factors using each x_i.
4. Identify all terms whose exponents sum to k. Each term's exponents form a composition.

If the parts have no restrictions, then we use the factor $x_i^1 + x_i^2 + x_i^3 + \ldots$ in step 2. Infinite sums and products are allowed. In practice, we simply ignore any terms above the range we are working in. For instance, to list all compositions of 5 into three parts, we multiply

$$(x_1^1 + x_1^2 + x_1^3 + \ldots)(x_2^1 + x_2^2 + x_2^3 + \ldots)(x_3^1 + x_3^2 + x_3^3 + \ldots)$$

and obtain the terms

$$x_1^1 x_2^2 x_3^2,\ x_1^2 x_2^1 x_3^2,\ x_1^2 x_2^2 x_3^1,\ x_1^3 x_2^1 x_3^1,\ x_1^1 x_2^3 x_3^1,\ x_1^1 x_2^1 x_2^3$$

with associated compositions (1, 2, 2), (2, 1, 2), (2, 2, 1), (3, 1, 1), (1, 3, 1), (1, 1, 3). Notice that terms with exponent 5 or higher never appear in the answer. Furthermore, the largest part needed is 3 since each part must be at least 1.

Exercise: Either count directly or use this method to calculate all compositions of 5. Where do the counts by number of parts appear in Pascal's triangle? What is the total number of compositions of 5?

Recall that the number of compositions of k into n parts is the combination $C(k-1, n-1)$. Thus, with n fixed, the generating function that counts compositions into n parts is

$$F(x) = x^n + nx^{n+1} + \ldots + C(k-1, n-1)x^k + \ldots.$$

On the other hand, in the products as in the algorithm above, to count all terms whose exponents sum to k, we set each x_i to the same x and record the coefficient of x^k. Simplifying, we have

$$F(x) = (x + x^2 + x^3 + \ldots)^n = \frac{x^n}{(1-x)^n}.$$

In the previous formula, the latter equality follows from the formal geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots.$$

If we compare both series for $F(x)$, we rediscover the familiar

Binomial Series: For $n \geq 0$, the coefficient of x^k in $1/(1-x)^n$ is the combination $C(k+n-1, k)$. In particular, for $n > 0$, the binomial series is the generating function that counts weak compositions into n parts.

Exercise: Write out the first four terms of the binomial series when $n = 1, 2, 3, 4$. For each n, where do the sequences of coefficients appear in Pascal's triangle?

Exercise: Find all weak compositions of 3 into n parts for $n = 1, 2, 3$, and verify using Pascal's triangle. Compare with weak compositions of 4 into three parts.

Exercise: Express the hockey stick rule as a counting rule for weak compositions. Interpret it based on choice of first part. How do we adjust the rule for strict compositions?

To obtain a generating function that counts with restricted parts, consider the case with parts 1 and 2. Let $p(x) = x + x^2$. If we extend the previous reasoning, the number of such compositions of k is the coefficient of x^k in the formal expression

$$F(x) = 1 + (x + x^2) + (x + x^2)^2 + (x + x^2)^3 + \ldots.$$

If we recast this sum as a geometric series, we have $F(x) = \dfrac{1}{1-x-x^2}$.

We used this series to deduce the Binet formula for the Fibonacci numbers!

Exercise: Use the binomial formula to express the Fibonacci numbers as sums of binomial coefficients. As a starting point for organizing coefficients, consider

$$\begin{array}{llllllllll} 1 & + & & & & & & & & \\ x & + & x^2 & + & & & & & & \\ x^2 & + & 2x^3 & + & x^4 & + & & & & \\ x^3 & + & 3x^4 & + & 3x^5 & + & x^6 & & + \ldots & \end{array}$$

For the left-justified Pascal's triangle, find the sums along the southwest-northeast diagonals.

Exercise: Explain the appearance of the Fibonacci numbers in Pascal's triangle by reconciling Pascal's identity with the Fibonacci recurrence relation. That is, explain why the sum over two adjacent diagonals in the left-justified Pascal's triangle is the sum over the next diagonal.

Exercise: Find the sum over the first k diagonals of the left-justified Pascal's triangle as an identity for partial sums $F_1 + F_2 + \ldots + F_k$ of Fibonacci numbers.

Exercise: Explain the previous sum directly as follows: consider a composition of length $k + 1$ with parts 1 and 2, and count by grouping compositions according to the maximum number of consecutive 1s at the end of the composition. For instance, such compositions end with exactly one of 2, 2 + 1, 2 + 1 + 1, …. What formula results if we organize the count by ending 2s?

In general, if $p(x)$ encodes the parts of interest, possibly with infinitely many terms, the generating function for compositions with these parts is given by

$$F(x) = 1/(1 - p(x)) = 1 + p(x) + p(x)^2 + p(x)^3 + \ldots.$$

If $p(x)$ has constant term zero, then each x^k appears in only finitely many terms in the sum. Thus, each coefficient is defined and obtained through finitely many operations.

If there are no restrictions on parts, then

$$p(x) = x + x^2 + x^3 + \ldots = x/(1 - x) \text{ and } F(x) = (1 - x)/(1 - 2x).$$

If we expand the geometric series, then the number of compositions of k is $2^k - 2^{k-1} = 2^{k-1}$.

Exercise: Verify the number of compositions of k by summing the appropriate row of Pascal's triangle. Interpret the terms of this sum.

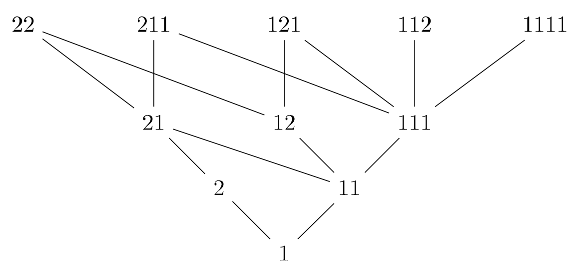
Exercise: Derive the closed form of the generating function for compositions restricted to parts with even values. That is, start with $p(x) = x^2 + x^4 + x^6 + \ldots = x^2/(1 - x^2)$. Then, substitute into the first equality for $F(x)$. How does this generating function differ from the generating function with no restrictions on parts? What happens to coefficients with odd index? Do the coefficients for even indices look familiar? Explain.

Exercise: Repeat the previous exercise with all odd parts. Does the generating function look familiar? Verify by listing compositions with odd parts for k = 1, 2, 3, 4, 5, 6. Describe the Fibonacci recursion by considering how to increase the length of a strip by 1 or 2 in this case.

Exercise: Repeat again with all parts allowed except 1. Verify that $p(x) = x^2/(1 - x)$ in this case. List the corresponding compositions for k = 1, 2, 3, 4, 5 and guess the closed formula. Describe the Fibonacci recursion by considering how to increase the length of a strip by 1 or 2 in this case.

Exercise: We have three types of compositions counted by the Fibonacci numbers: compositions into parts of size 1 or 2, into parts bigger than 1, and into parts that are odd. Can you give a bijective proof of these matchings? That is, when the counts agree, can you pair up the elements of one set of compositions to the elements of the others without reference to Fibonacci numbers or generating functions? (Hint: Every integer greater than 1 is a sum of the form 2 + 1 + … + 1. How do we ensure every composition starts with a 2? Likewise, every positive odd integer is a sum of the form 1 + 2 + … + 2. How do we ensure every composition starts with a 1?)

Exercise: For more insight on the hint in the previous exercise, consider the Hasse diagram for the graded poset of compositions with parts 1 and 2. The partial ordering is given by entry-wise comparison, starting at the left-hand side; we assume hidden 0s on the right when the number of parts differ. Use the diagram below to construct the other two cases; the hint suggests we precede all entries by either a 1 or 2 and then translate (by using the pairing from the previous exercise). The pairing induces a partial order on compositions into parts greater than 1 that differs from the standard partial order on compositions. For example, in the induced order, $5 \leq 3 + 3$ but $2 + 2$ is not less than $3 + 3$. Describe the up linear operators in each case.



What lies beyond the Fibonacci numbers? For a glimpse, we restrict to parts 1, 2, and 3 and imitate the derivation of the Binet formula for Fibonacci numbers. Please refer to Volume 15 Issue 6 for the earlier derivation; here, we outline only the major changes for the new model. In particular, we avoid a long digression by approximating roots instead of deriving them exactly.

Exercise: Suppose $p(x) = x + x^2 + x^3$. Use the expanded formula with $p(x)$ and the trinomial formula to compute the number of compositions of k = 1, 2, 3, 4 with parts 1, 2, and 3. Verify by listing the compositions explicitly, and draw the corresponding strips.

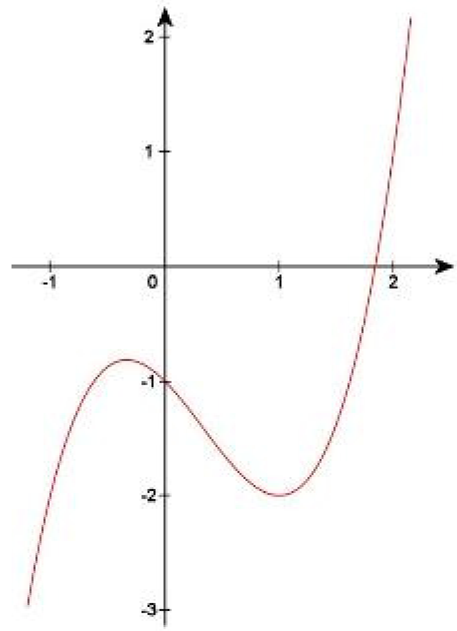
Definition. The tribonacci numbers are defined by $T_0 = T_1 = 0$, $T_2 = 1$, and

$$T_k = T_{k-1} + T_{k-2} + T_{k-3}, \text{ for } k > 2.$$

Exercise: To derive an analogue of the Binet formula, assume T_k has the form

$$T_k = A\phi_0^k + B\phi_1^k + C\phi_2^k.$$

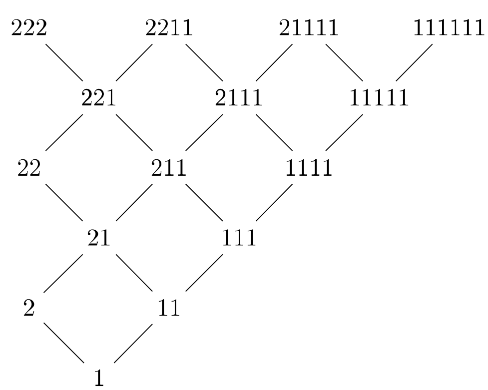
- Calculate T_k up to $k = 20$ using the definition. Then calculate the ratios T_k/T_{k-1}. Let ϕ_0 be the limit of these ratios. With $k = 20$, we get the first six decimal places of ϕ_0.
- In the graph of $f(x) = x^3 - x^2 - x - 1$ at right, the single real zero is ϕ_0. Verify by evaluating $f(x)$ at your approximation for ϕ_0.
- Use synthetic division and the quadratic formula to approximate the size of the complex roots ϕ_1 and ϕ_2. In our formula above, their contributions tend to zero quickly for large k.
- With a good approximation for ϕ_0, the value of T_{k+1} is obtained by multiplying T_k by ϕ_0 and rounding to the nearest integer. Calculate up to T_{25} from T_{20} this way and verify with the recursive rule.
- Use your data set to estimate A in the approximation $T_k \sim A\phi_0^k$, approximate T_{25}, and compare with the actual value.
- Verify that the numbers of compositions of k in the previous exercise satisfy the tribonacci recurrence relation. Describe these counts in terms of the tribonacci numbers.



Finally, where are the Catalan numbers? If we construct the poset of compositions with n parts and with parts restricted to 1 and 2, we obtain the Boolean posets from the previous installment. If we remove the ordering condition on compositions, we obtain instead **partitions** of k into n parts. In this case, it is customary to write the sum in nonincreasing order. For instance, the partitions of 5 are given by

$$5, \quad 4+1, \quad 3+2, \quad 3+1+1, \quad 2+2+1, \quad 2+1+1+1, \quad 1+1+1+1+1.$$

There is plenty to say about partitions. For now, we consider the partitions of k with parts 1 and 2. We define a graded poset for these partitions using the same partial ordering as in the exercise above. With only parts 1 and 2, there are at most two ways to cover a partition: by turning the first 1 into a 2, or by adding a 1 to the end. The Hasse diagram of such a poset begins:



The full Hasse diagram of these partitions has the shape of the Catalan triangle, from which we obtain Catalan numbers by counting Dyck paths, or the chains from 1 to the entries in the leftmost column.

The Needell in the Haystack[1]

Natural Language Processing, Part 2: Language Models
by Anna Ma | edited by Jennifer Sidney

In this installment of "The Needell in the Haystack," we continue discussing Natural Language Processing (NLP). NLP is an area of artificial intelligence that lives at the cross-section of Linguistics (the study of language and its structure) and Computer Science. In NLP, the goal is to train a computer to understand, interpret, and generate human text. In the last "The Needell in the Haystack" installment, we discussed a simple model for auto-completing the ends of sentences to answer questions such as the following:

Given the sentence

"*I think math is very* _____ ."

what should go into the blank space? We showed that with a minimal toy data set, we can train and implement a method that auto-completes this sentence. In this installment of "The Needell in the Haystack," we will explore ways to quantify the likelihood of *sequences of words* using *language models*.

Transition Matrices

Here's a quick recap of what we discussed in the last installment of "The Needell in the Haystack." First, we started with a small toy data set:

	think	math	fun	mango	squishy	challenging	problem	rewarding
think	0	2/2	0	0	0	0	0	0
math	0	0	3/4	0	0	1/4	0	0
fun	0	0	0	0	1/3	0	1/3	1/3
mango	0	0	0	0	2/2	0	0	0
squishy	0	0	0	1/1	0	0	0	0
challenging	0	0	1/2	1/2	0	0	0	0
problem	0	0	0	0	0	0	0	0
rewarding	0	0	0	0	0	0	0	0

Table 1. Transition matrix for toy data set.

"*I think math is very fun. I think math is fun. My mango is squishy. Math has challenging and fun problems. It is challenging if my mango is not squishy. I find math very fun and rewarding. I have fun when I have my squishy mango.*"[2]

After stop words (words such as "I," "is," "my," "the") were removed, we used co-occurrences (word pairs)

"*think math // math fun // think math // math fun // mango squishy // Math challenging // challenging fun // fun problems // challenging mango // mango squishy // math fun // fun rewarding // fun squishy // squishy mango.*"

[1] This content is supported in part by a grant from MathWorks. Anna Ma is an Assistant Professor at the University of California Irvine.
[2] Anna was gifted a squishy mango-shaped stress ball, and she (still) loves it. It lives in her office and has helped on many stressful occasions!

to construct a **transition matrix** as shown in Table 1 above.

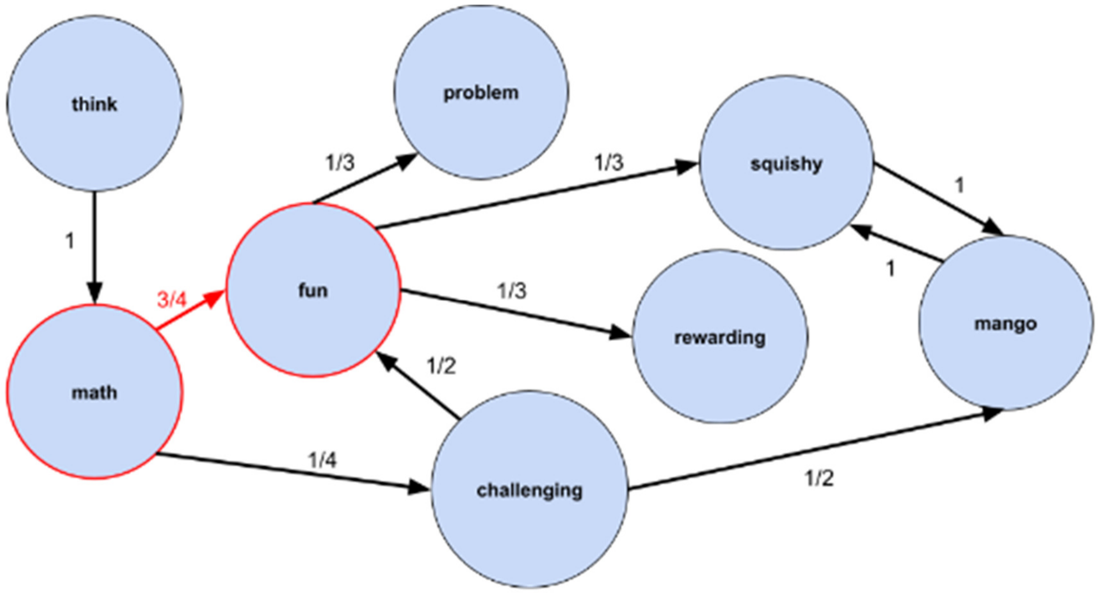


Figure 1. Visualization of the transition matrix *M*.

We interpret the transition matrix M as follows: M_{ij} represents the probability that word j follows next after word i. For example, the word j = "*math*" follows the word i = "*think*" with probability 1 because in the "*think*" row and "*math*" column, the matrix value is $M_{ij} = 1$. The transition matrix can be visualized as a **directed graph** as shown in Figure 1. In the graph, each circle (node) represents a word, and each directed arrow (edge) points to what the next word can potentially be. The values next to the arrows correspond to the probability that the word being pointed to is the next word.

We use a **maximum likelihood** approach to determine the next word in our simple autocomplete model. For example, if we look at the "*math*" row, the word "*fun*" has a 3/4 chance to appear next. In contrast, the word "*challenging*" has a 1/4 chance of appearing next. Thus, the autocomplete would finish the sentence "I think math is very _____" with the word "*fun*," because "*fun*" has the largest probability of occurring next. In Figure 1, we can look at the node corresponding to the word "*math*" and see which edge has a larger number – in this case, "*fun*," as highlighted in red.

Language Models and Probability

Language models are probability distributions over words; they assign likelihoods to sequences of words. Not all sequences of words are likely to appear together. More importantly, the order in which words appear matters. Consider the following two sequences of words:

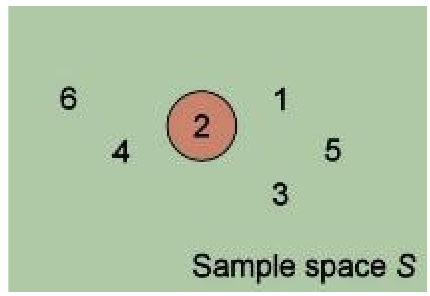
Sequence S1: "*Jumps over lazy a brown fox quick the dog.*"
Sequence S2: "*The quick brown fox jumps over a lazy dog.*"[3]

[3] This sentence is a **pangram**, a sentence that contains all letters of the English alphabet.

What is the likelihood, or probability, that sequence of words S1 appears in a textbook? What is the probability sequence of words S2 appears? Since S1 is not a coherent sentence, we might conjecture that the probability $P(\text{S1})$ of sequence S1 occurring is 0, whereas $P(\text{S2}) > 0$. In what follows, we will use concepts in probability to quantify the likelihood that words appear together in a particular order.

Before moving on, let's first lay out some groundwork in probability. **Probability Theory** is a branch of mathematics that is concerned with the analysis of random events. A **probability** is a measure, on a scale of 0 to 1, of how likely an event is to occur. Formally, a probability is the output of a function P defined on the subsets of a set S called the **sample space**, which represents the set of all possible outcomes. The function P maps subsets of the sample space, called **events**, to a real number between 0 and 1, which gives the probability of that event occurring. Let's now look at a few examples.

Example 1a. Suppose a fair 6-sided die is rolled. The sample space is $S = \{1, 2, 3, 4, 5, 6\}$ and the probability of rolling, say, a 2, is 1/6, so $P(\{2\}) = 1/6$. The figure at right shows the sample space (green) and the outcome of interest (red). In this example, the probability that an even number is rolled is 3/6; that is, $P(\{2, 4, 6\}) = 3/6$.



Example 2a. Let the sample space S be the sample space of all possible word pairs in our toy data set, and let the probability of selecting a word pair be proportional to the number of times that the co-occurrence appears in the toy data set. For example, the probability for "*math fun*" is 3/14, as there are 14 co-occurrences and "*math fun*" appears 3 times. As another example, the probability that we pick a co-occurrence that starts with the word "*math*" is 4/14, since there are 4 co-occurrences that start with the word "*math*" ("*math fun*" appears 3 times, and "*math challenging*" appears once).

In most languages, the order in which words appear matters, so the likelihood a word next appears depends on the previous word. This structure can be captured using **conditional probabilities**. Conditional probabilities measure the likelihood of an event E occurring given the additional information that the outcome was among a certain subset of outcomes C. The subset C of S is referred to as the **condition**. We denote this conditional probability by $P(E \mid C)$, which is read, "the probability that E occurs given C." We can compute $P(E \mid C)$ using the formula

$$P(E \mid C) = \frac{P(E \cap C)}{P(C)}. \tag{1}$$

Let's try to get a better understanding of this equation. The probability of E given C is the probability that an outcome is in E, given that we know that the outcome is in C. Imagine doing the random experiment (such as rolling a die) over and over – n times – for some huge number n. $P(E \mid C)$ is supposed to tell us what fraction we would expect of the outcomes that are in C are also in E. We expect a fraction of $P(C)$ of the outcomes to be in C, that is, $nP(C)$ of the experiments will have an outcome in C. Of these, we expect $n\,P(E \cap C)$ of the outcomes to be in both E and C. So the fraction of outcomes that come out in C that are also in E is $n\,P(E \cap C) / nP(C)$, which simplifies the right-hand side of Equation 1.

Another way to interpret conditional probability is to view it as a change in sample space. For example, suppose a fair 6-sided die is rolled. Given that an even number was rolled, what is the probability that a 2 was rolled? Since we know the roll has to be an even number, there are three possible values that could have been rolled: 2, 4, or 6. Since each possible roll is equally likely, $P(\{2\} \mid \{2, 4, 6\}) = 1/3$. In effect, we are changing the sample space from $\{1, 2, 3, 4, 5, 6\}$ to $\{2, 4, 6\}$. We can think of $P(E \cap C)$ as a measure of relative likelihood of events given that the outcome is in C, and we can think of $P(C)$ as a **renormalization** factor to adjust these likelihoods back to probabilities.

Example 1b. Suppose two fair 6-sided dice are rolled. Given that the maximum number rolled was less than 5, what is the probability that the sum of the numbers on the two dice is 7? For the rolling of two dice, the sample space consists of the 36 pairs (a, b) where a and b are both in the set $\{1, 2, 3, 4, 5, 6\}$. The probability of any particular pair coming up is 1/36 (since the dice are fair). Here, we are asked to compute the conditional probability of E given C where E is the set of pairs (a, b) in the sample space where $a + b = 7$ and C is the set of pairs (a, b) in the sample space where $\max(a, b) < 5$. Now $P(C) = 16/36$ and $P(E \cap C) = 2/36$. Therefore, $P(E \mid C) = (2/36)/(16/36) = 1/8$.

1, 1	1, 2	1, 3	1, 4	1, 5	**1, 6**
2, 1	2, 2	2, 3	2, 4	**2, 5**	2, 6
3, 1	3, 2	3, 3	**3, 4**	3, 5	3, 6
4, 1	4, 2	**4, 3**	4, 4	4, 5	4, 6
5, 1	**5, 2**	5, 3	5, 4	5, 5	5, 6
6, 1	6, 2	6, 3	6, 4	6, 5	6, 6

Sample space for rolling two dice. Outcomes in C are shaded blue. Boldfaced outcomes constitute E. Red outcomes are in E and C.

Example 2b. What is the probability that "*fun*" is the next word in a sequence of words when "*math*" is the preceding word? Using Equation 1,

$$P(\text{“fun” is next} \mid \text{“math” is preceding}) = \frac{P(\text{"}math\ fun\text{"})}{P(\text{"}math\text{" is preceding})} = \frac{3/14}{4/14} = 3/4\,.$$

(Note that the event "*math fun*" consists of the outcomes where "*fun*" is next *and* "*math*" is preceding.) This is how the probabilities next to the arrows in Figure 1 are computed. Those numbers *are* conditional probabilities. The number next to an arrow is the conditional probability that the next word is the word the arrow is pointing to, given that the preceding word is the word that the arrow is pointing from. Thus, the number next to the arrow pointing from "*math*" to "*fun*" is 3/4. Can you check that the number next to the arrow pointing from "*math*" to "*challenging*," 1/4, is the correct conditional probability

$$P(\text{“challenging” is next} \mid \text{“math” is preceding})?$$

Example 3. We can also use conditional probabilities to compute the probability of a sequence of more than two words. For simplicity, in the remainder of this article, we will drop the "is next" and "is preceding" in our notation and interpret the conditional probability $P(\text{"}fun\text{"} \mid \text{"}math\text{"})$ as "the probability *fun* immediately follows *math*." By rearranging Equation 1, we see that $P(E \cap C) = P(C)P(E \mid C)$. Now, if we want to compute the probability for a sequence of more than two words, for example, $P(\text{"}math\ fun\ problems\text{"})$ we can write

$$P(\text{"}math\ fun\ problems\text{"}) = P(\text{"}math\text{"})P(\text{"}fun\ problems\text{"} \mid \text{"}math\text{"}).$$

We can also write

$$P(\text{"}fun\ problems\text{"} \mid \text{"}math\text{"}) = P(\text{"}fun\text{"} \mid \text{"}math\text{"})P(\text{"}problems\text{"} \mid \text{"}math\ fun\text{"}).$$

And under the assumption that the next word only depends on the preceding one,

$$P(\text{"}problems\text{"} \mid \text{"}math\ fun\text{"}) = P(\text{"}problems\text{"} \mid \text{"}fun\text{"}).$$

Let's put this all together:

$$P(\text{"}math\ fun\ problems\text{"}) = P(\text{"}math\text{"})P(\text{"}fun\text{"} \mid \text{"}math\text{"})P(\text{"}problems\text{"} \mid \text{"}fun\text{"}).$$

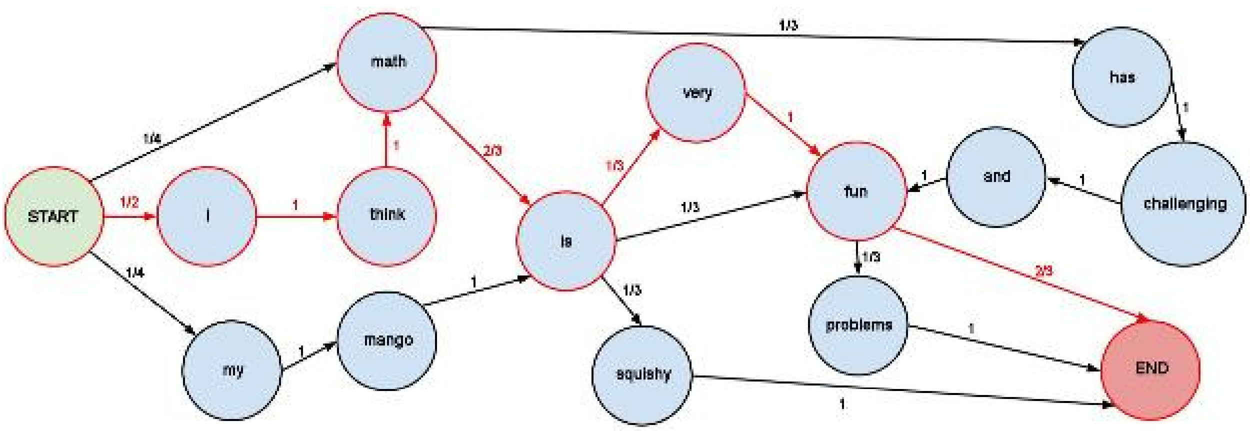


Figure 2. Visualization of the transition matrix for the first four sentences, including stop words and START/END sentence indicators.

Generating Sentences

What if we want to string together a sequence of words? How would we compute the probability that a particular sequence occurs? Is it even possible to use the simple transition matrix method for writing entire sentences? Suppose we start a sentence with the phrase "*I think* …" and try to use our transition matrix to complete this sentence with more than one word.[4] Looking at

[4] Since "problem," "challenge," and "rewarding" have the same probability, we pick any one of these options.

Figure 1 and choosing the most likely words, we get "*I think math* …." Continuing this process, we might end up with

"*I think ...*" → "*I think math ...*" → "*I think math fun ...*" → "*I think math fun problem*,"

which doesn't seem to make a lot of sense. Perhaps it means "*I think math has lots of fun problems*," or maybe "*I think math is fun because of the problems I do*." There's a lot of ambiguity without stop words. One way to fix this is to include the stop words in our transition matrix.

Figure 2 shows the directed graph for the first four sentences in the toy data set where all words are included, and we use START/END indicators for the beginning and end of each sentence. In red, we've highlighted a possible completion for "*I think* …" using maximum likelihood to choose the next word. The sentence completion is "*I think math is very fun*," which is a sentence that appears in our data set!

A language model is a function P that maps sequences of words that start with START and end with END to a number between 0 and 1. The probability for each sequence can be computed using conditional probabilities. In the most basic language model, we assume that the next word only depends on the proceeding word. For example,

$$
\begin{aligned}
&P(\text{START "}\textit{I think math is very fun}\text{" END}) \\
&\quad = P(\text{START})P(\text{"}\textit{I}\text{" | START}) \cdot P(\text{"}\textit{think}\text{" | "}\textit{I}\text{"}) \cdot P(\text{"}\textit{math}\text{" | "}\textit{think}\text{"}) \cdot P(\text{"}\textit{very}\text{" | "}\textit{math}\text{"}) \\
&\qquad \cdot P(\text{"}\textit{fun}\text{" | "}\textit{very}\text{"}) \cdot P(\text{END | "}\textit{fun}\text{"}) \\
&\quad = 1 \cdot 1/2 \cdot 1 \cdot 1 \cdot 2/3 \cdot 1/3 \cdot 1 \cdot 2/3 = 4/54.
\end{aligned}
$$

If we move away from the maximum likelihood approach, we can form new sentences and compute their likelihood of occurring. For instance, we can form the sentence "*My mango is very fun*," which is not a sentence that appears in the toy data set, by following the path from START to END in Figure 2 and passing through each of those words.

Contextual Awareness

You may have noticed that not all sequences of words will make sense; for example, we can form the sentence "*I think math is squishy*" with our current data. Let's consider the sentence "*My favorite subjects in school are math* _____ ." How would the language model complete this sentence? The deficiencies in this approach stem from a few sources. First, there's insufficient training data. Note that our data set doesn't contain any information about different kinds of subjects. Math is the only subject that appears in the set of possible words that the matrix M can provide. Second, this approach lacks contextual awareness of the sentence. No matter what was conveyed earlier, we only look at the most recent word to determine the next word. Stay tuned for the next installment of "The Needell in the Haystack," where we will discuss improving the current language model to include contextual awareness!

Romping Through the Rationals, Part 4

by Ken Fan | edited by Jennifer Sidney

Emily and Jasmine are studying the sequence obtained by starting with 0, then successively applying the function $f(x) = 1/(1 + \lfloor x \rfloor - \{x\})$ to its terms. A social media post Jasmine read claims that every nonnegative rational number occurs in the sequence exactly once.

Jasmine: What is it about this banana bonanza that's so great?

Emily: It's the pervasion of banana flavor in an already mouthwatering, cake-like bread.

Emily and Jasmine each shave off a forkful of banana bonanza from opposite ends of the loaf, take it into their mouths, and treat their tongues to the tantalizing texture and taste. Suddenly, Jasmine makes an excited, muffled sound and hurriedly swallows.

Jasmine: I just got an idea!

Emily: Mmm … hmm? What idea?

Jasmine: When we were thinking about integer sequences whose ratios of consecutive terms romp through the nonnegative rational numbers, we found that it wasn't hard to create a sequence that never repeated a rational number. But we were stuck on how to ensure that it produces *every* nonnegative rational number.

Emily: Yes?

Jasmine: We tried to think of a way to modify a sequence whose ratios of consecutive terms produced an incomplete list of rational numbers so that it would produce more, but we came up empty.

Emily: You thought of a way?

Jasmine: Oh, no. I didn't. But now we know two such sequences: the one we created, and the one formed by the numerators of the sequence we learned from the internet. It made me wonder, can we morph one sequence into the other? And that made me think whether – instead of trying to create a sequence from scratch – maybe we can get other sequences by modifying existing ones.

Emily: I see! Instead of finding a way to add in missing rational numbers, you want to focus on ways to reorder them.

Jasmine: Yes!

Emily: We'd need some kind of morphing operation.

Jasmine: Right, and that's the idea I just had! I know a way to modify such sequences to get new ones. Let's call a sequence a_k of nonnegative integers a "rational romper" if consecutive terms are relatively prime and the ratios a_k/a_{k+1} yield each of the nonnegative rational numbers exactly once.

Emily: Okay. I like that name because it reflects the fact that any list of all the nonnegative rational numbers must jump around.

Jasmine: Suppose a_k is a rational romper, three of its consecutive terms are A, B, and C, and A and C are relatively prime. Because the sequence is a rational romper, the 2-term subsequence A, C must occur somewhere in the sequence since we must be able to get the rational number A/C as a ratio of consecutive terms. That means that we can modify the sequence by removing the B from the subsequence A, B, C and placing it in between the A and C in the 2-term subsequence A, C!

Emily: I see! In fact, you can do this starting with any subsequence $s_1, s_2, s_3, \ldots, s_n$, where $n > 2$ and s_1 is relatively prime to s_n. It doesn't have to be a 3-term subsequence. We know that s_n must appear right after s_1 somewhere in the sequence; if it doesn't occur within our subsequence $s_1, s_2, s_3, \ldots, s_n$, we can remove the middle terms $s_2, s_3, \ldots, s_{n-1}$ from $s_1, s_2, s_3, \ldots, s_n$ and insert it back into the sequence between the consecutive occurrences of s_1 and s_n.

Jasmine: Yes, you're right! If we change the sequence

$$\ldots, s_1, s_n, \ldots, s_1, s_2, s_3, \ldots, s_{n-1}, s_n, \ldots$$

to

$$\ldots, s_1, s_2, s_3, \ldots, s_{n-1}, s_n, \ldots, s_1, s_n, \ldots,$$

every ratio of consecutive terms that appears in the first will also appear in the second. In the second sequence, the ratio s_1/s_n will come later, whereas the ratios s_i/s_{i+1}, $1 \le i \le n-1$ will still appear consecutively, but earlier. All the other ratios will appear in the same relative order. And an analogous picture holds if the 2-term subsequence s_1, s_n appears after the subsequence $s_1, s_2, s_3, \ldots, s_n$ instead of before. In fact, it just reverses the picture.

Emily: It's sequence splicing …

Jasmine: Okay, let's call it splicing! Do you think this splicing tool will provide enough flexibility that any rational romper can be changed into any other?

Emily: I don't know, but it sure seems like you could modify a rational romper a whole lot with this technique!

Jasmine: Let's see if we can modify any rational romper using this subsequence relocation technique so that it begins 0, 1, 1.

Emily: Good idea! All rational rompers begin 0, 1, so you're wondering if we can modify any rational romper so that the third term is a 1. And, yes, I think we can! We know that the 2-term subsequence 1, 1 must occur somewhere in the rational romper because it must produce the number 1 as the ratio of some two consecutive terms. If the third term is *not* 1, then we find wherever the second occurrence of 1 is. Let's say it is the term a_n. So the sequence goes

$$0, 1, a_3, a_4, \ldots, a_{n-1}, 1, \ldots.$$

We know that none of the terms $a_3, a_4, \ldots, a_{n-1}$ are equal to 1, by construction, so the 2-term subsequence 1, 1 doesn't occur among the first n terms. It has to occur later – possibly with the nth and $(n + 1)$th terms, possibly thereafter. Either way, we can perform a splice, moving the subsequence $a_3, a_4, \ldots, a_{n-1}$ to in between the two consecutive ones and bringing the first two occurrences of 1 next to each other!

Jasmine: That works out so nicely. I think almost the same argument allows us to modify any rational romper so that it begins 0, 1, p, for any positive integer p. You just showed what to do if $p = 1$. If $p > 1$, we look for the first occurrence of p. Again, say it is the nth term so $a_n = p$ and a_k is not p for any $k < n$. If $p = 3$, the sequence already begins 0, 1, p, so assume $p > 3$. We know that the 2-term subsequence 1, p must occur somewhere in the sequence because some two consecutive terms have ratio $1/p$, and this 2-term subsequence … hmm … it may be slightly more complicated, because the 2-term subsequence 1, p might occur with the $(n - 1)$th and nth terms, that is, there might be a 1 just before the first p. Or it could occur after the nth term.

Emily: If it occurs after the nth term, there's no problem; we can remove the subsequence $a_3, a_4, \ldots, a_{n-1}$ from between the first 1 and the first p to between the 1 and p of the 2-term subsequence 1, p. So what happens if $a_{n-1} = 1$? Then our sequence looks like
$0, 1, a_3, a_4, \ldots, a_{n-2}, 1, p, \ldots.$

Jasmine: Unfortunately, we can't simply move the p and place it between the 1 and a_3, because the term after p might not be equal to a_3. That would mean the sequence would no longer render the rational number p/a_{n+1}, and it would render p/a_3 twice.

Emily: You're right: the term after that first p may not be a_3 … but if it's not, we can look for the first term after that first p which *is* equal to a_3. Let's say m is the smallest positive integer greater than n for which $a_m = a_3$. So then our sequence looks like

$$0, 1, a_3, a_4, \ldots, a_{n-2}, 1, p, \ldots, a_{m-1}, a_m = a_3, \ldots.$$

Jasmine: Great idea! Then we can remove the subsequence $p, a_{n+1}, \ldots, a_{m-1}$, which is bookended by 1 on the left and $a_m = a_3$ on the right, and we can insert it between the first 1 and a_3! That would produce the sequence

$$0, 1, p, \ldots, a_{m-1}, a_3, a_4, \ldots, a_{n-2}, 1, a_m = a_3, \ldots,$$

which *does* begin 0, 1, p! This is exciting!

Emily: What else can we do with this splicing technique?

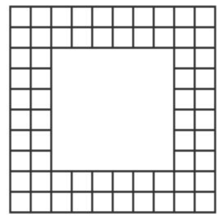
To be continued …

Domino Tilings of a Square Annulus
by Ken Fan

For this investigation, Charlotte and Kathy worked primarily with Girls' Angle mentors Laura Pierson and Valerie Yuen.

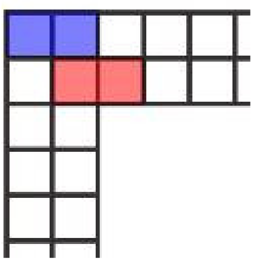
Tiling is a topic that offers a wealth of fascinating mathematical questions. In the last issue, Addie Summer counted the number of ways to tile a certain zigzag path using dominos. Here, we follow the thoughts of Girls' Angle members Kathy Lin and Charlotte Younger as they count the number of ways to tile a square annulus using dominos, and then Kathy, who went on to prove a beautiful property of the resulting sequence.

Let n be a nonnegative integer. Consider an $n + 4$ by $n + 4$ grid of squares. Now, remove the middle n-by-n grid. When n is positive, the result is a square annulus and when $n = 0$, the result is a 4 by 4 grid, which we can think of as an annulus in the same family as the others, but without a hole. How many ways are there to tile these annuli with 1 by 2 dominos?

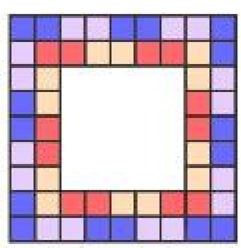


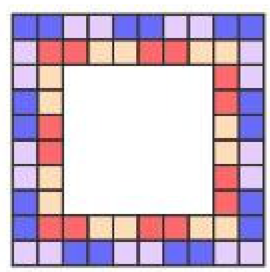
Annulus corresponding to $n = 6$

The corner squares of the annulus drew Charlotte and Kathy's attention, because those squares can be covered by a domino in only two ways, whereas all the other squares can be covered in more than two ways. For this reason, they decided to consider $2^4 = 16$ cases, corresponding to each of the 16 ways in which the four corner squares may be covered with a domino.



However, before considering these 16 cases, they made the following simplifying observation. Suppose that a corner of the annulus is tiled as shown at left. It turns out that there is one and only one way to complete the tiling of the annulus. This can be checked directly by completing the tiling and noting that as tiles are placed, there is always a square which can be tiled by a domino in only one way. See below for illustrations of the completions to tilings of the annuli $n = 5$ and $n = 6$.

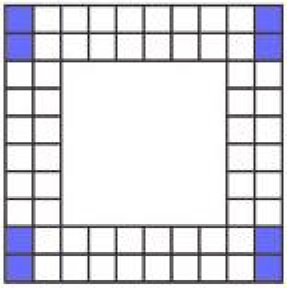




By symmetry, there is also one and only one way to complete a tiling where the upper left corner square is covered with a vertical domino, and a second domino is placed to its right and shifted down one square (illustrated by flipping the diagram above over the northwest-southeast diagonal). The analysis of the 16 cases is simplified if we ignore these two tilings and add them back in later. Let's refer to these two special tilings as **staggered tilings.**

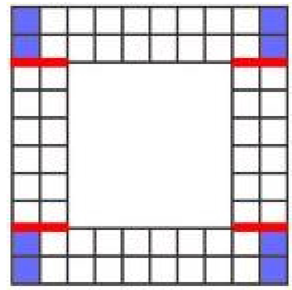
Also, note that flipping the annulus around the northwest-southeast diagonal is a symmmetry of the annulus that swaps tilings where the upper left square is covered by a vertically-oriented domino with tilings where the upper left square is covered by a horizontally-oriented domino. Hence, we can count all the tilings by focusing attention on those tilings where the upper left square is covered by a vertically-oriented domino, then doubling the number of such tilings. This reduces the number of cases we have to consider from 16 to 8.

Let’s now follow their detailed consideration of these 8 cases. For the first of these 8 cases, consider the one where all four corner squares are covered with vertically-oriented dominos.



Because we are ignoring staggered tilings, we may assume that no domino will straddle any of the red lines in the figure at right.

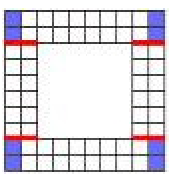
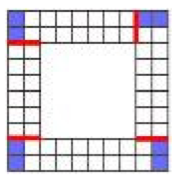
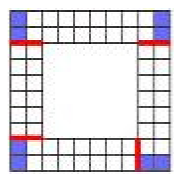
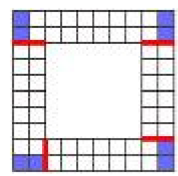
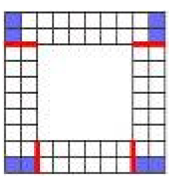
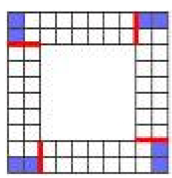
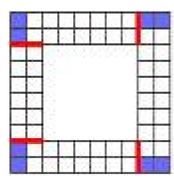
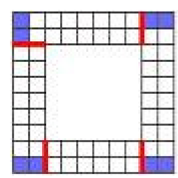
This splits the untiled parts of the annulus into four rectangular subsections, each of which may be independently tiled with dominos.



It is a well-known fact that a $2 \times m$ rectangle may be tiled by dominos in exactly F_{m+1} ways, where F_k is the Fibonacci sequence which begins $F_1 = F_2 = 1$, and satisfies $F_{n+1} = F_n + F_{n-1}$ for $n > 1$. (If you’re unfamiliar with this fact, try to prove it using induction.)

Therefore, the number of tilings covered by this case is equal to $F^2_{n+1}F^2_{n+3}$.

All 8 cases follow a similar pattern, summarized in the table below.

 $F^2_{n+1}F^2_{n+3}$	 $F_{n+1}F^2_{n+2}F_{n+3}$	 $F_{n+1}F^2_{n+2}F_{n+3}$	 $F_{n+1}F^2_{n+2}F_{n+3}$
 $F_{n+1}F^2_{n+2}F_{n+3}$	 F^4_{n+2}	 $F_{n+1}F^2_{n+2}F_{n+3}$	 $F_{n+1}F^2_{n+2}F_{n+3}$

Notice that six of the case counts are the same. In total, these 8 cases give us

$$F^2_{n+1}F^2_{n+3} + 6F_{n+1}F^2_{n+2}F_{n+3} + F^4_{n+2}$$

tilings. We double this to include the tilings where the upper left square is covered by a horizontally-oriented domino, and then add the two staggered tilings to get the total number of tilings of the square annulus with an n-by-n hole: $2(F^2_{n+1}F^2_{n+3} + 6F_{n+1}F^2_{n+2}F_{n+3} + F^4_{n+2} + 1)$.

Here are the first nine terms of this sequence, starting with $n = 0$:

36, 196, 1444, 9604, 66564, 454276, 3118756, 21362884, 146458404, …

Do you notice anything curious about these numbers?

They’re all prefect squares!

$$6^2, 14^2, 38^2, 98^2, 258^2, 674^2, 1766^2, 4622^2, 12102^2, \ldots$$

Are all the counts of the numbers of tilings of these square annuli perfect squares?

Here’s Kathy:

We use the well-known identity $F_{n-1}F_{n+1} = F_n^2 + (-1)^n$, and proceed as follows:

$$\begin{aligned}
&2(F_{n+1}^2F_{n+3}^2 + 6F_{n+1}F_{n+2}^2F_{n+3} + F_{n+2}^4 + 1) \\
&\quad = 2((F_{n+2}^2 + (-1)^{n+2})^2 + 6F_{n+2}^2(F_{n+2}^2 + (-1)^{n+2}) + F_{n+2}^4 + 1) \\
&\quad = 2(F_{n+2}^4 + 2F_{n+2}^2(-1)^{n+2} + 1 + 6F_{n+2}^2(F_{n+2}^2 + (-1)^{n+2}) + F_{n+2}^4 + 1) \\
&\quad = 2(F_{n+2}^2(F_{n+2}^2 + 2(-1)^{n+2} + 6(F_{n+2}^2 + (-1)^{n+2}) + F_{n+2}^2) + 2) \\
&\quad = 2(F_{n+2}^2(8F_{n+2}^2 + 8(-1)^{n+2}) + 2) \\
&\quad = 4(4F_{n+2}^2(F_{n+2}^2 + (-1)^{n+2}) + 1) \\
&\quad = 4(4F_{n+2}^4 + 4F_{n+2}^2(-1)^{n+2} + 1) \\
&\quad = 4(2F_{n+2}^2 + (-1)^{n+2})^2 \\
&\quad = 4(F_{n+2}^2 + F_{n+2}^2 + (-1)^{n+2})^2 \\
&\quad = 4(F_{n+2}^2 + F_{n+1}F_{n+3})^2 \\
&\quad = 4(F_{n+2}^2 + F_{n+1}^2 + F_{n+1}F_{n+2})^2 \\
&\quad = (2F_{n+2}^2 + 2F_{n+1}^2 + 2F_{n+1}F_{n+2})^2 \\
&\quad = (F_{n+2}^2 + F_{n+1}^2 + F_{n+2}^2 + 2F_{n+1}F_{n+2} + F_{n+1}^2)^2 \\
&\quad = (F_{n+2}^2 + F_{n+1}^2 + (F_{n+2} + F_{n+1})^2)^2 \\
&\quad = (F_{n+2}^2 + F_{n+1}^2 + F_{n+3}^2)^2 \\
&\quad = (F_{n+1}^2 + F_{n+2}^2 + F_{n+3}^2)^2 .
\end{aligned}$$

Thus, we see that the number of ways to tile the annulus made by removing an n-by-n hole from the middle of an $n + 4$ by $n + 4$ grid of squares is, indeed, a perfect square, and that, in fact, they are the squares of the sum of the squares of three consecutive Fibonacci numbers!

Editor’s note: The results Charlotte and Kathy found are a rediscovery of results of Tauraso, published in his paper “A New Domino Tiling Sequence,” which appeared in the Journal of Integer Sequences, Volume 7, 2004 as Article 04.2.3. We think it notable that they found these results independently less than 20 years after Tauraso published his paper.

The sequence that counts the number of tilings of the annulus appears as sequence A220436 in the On-Line Encyclopedia of Integer Sequences. The sequence of its square roots appears there as sequence A127546.

These sequences relate to Candido’s identity, which states that

$$(x^2 + y^2 + (x + y)^2)^2 = 2(x^4 + y^4 + (x + y)^4).$$

If you substitute $x = F_n$, $y = F_{n+1}$, one sees that the number of ways to tiling the annuli that Charlotte and Kathy considered is also given by twice the sum of the 4th powers of three consecutive Fibonacci numbers.

Notes from the Club

These notes cover some of what happened at Girls' Angle meets. In these notes, we include some of the things that you can try or think about at home or with friends. We also include some highlights and some elaborations on meet material. Less than 5% of what happens at the club is revealed here.

Session 32 - Meet 5 March 2, 2023	Mentors:	Elisabeth Bullock, Jade Buckwalter, Anushree Gupta, Clarise Han, Hanna Mularczyk, Kate Pearce, AnaMaria Perez, Laura Pierson, Vievie Romanelli, Swathi Senthil, Jane Wang, Jing Wang, Rebecca Whitman, Muskan Yadav

Some members made crocheted models of the hyperbolic plane (à la the work of Daina Taimina who gave us a wonderful presentation about such things on March 4, 2021). These models included a pentagonal tiling and an illustration of the fact that in the hyperbolic plane, through a point outside a line, there are multiple lines that do not intersect the given line (contrary to Euclid's fifth postulate).

Session 32 - Meet 6 March 9, 2023	Mentors:	Elisabeth Bullock, Anne Gvozdjak, Clarise Han, Padmasini Krishnan, Hanna Mularczyk, Kate Pearce, AnaMaria Perez, Laura Pierson, Vievie Romanelli, Swathi Senthil, Jane Wang, Jing Wang, Rebecca Whitman, Muskan Yadav, Valerie Yuen

Some members embarked on a project of finding an equation whose graph produces a Valentine heart shape. The members eventually settled on a heart shape composed out of four circular arcs, each tangent to two others.

Session 32 - Meet 7 March 16, 2023	Mentors:	Elisabeth Bullock, Jade Buckwalter, Anne Gvozdjak, Clarise Han, Padmasini Krishnan, AnaMaria Perez, Laura Pierson, Vievie Romanelli, Jane Wang, Jing Wang, Rebecca Whitman, Muskan Yadav, Valerie Yuen

Some members made further exploration of a coordinate system on the plane based on the distances of each point from 3 intersecting lines. This is related to barycentric coordinates.

Session 33 - Meet 8 March 23, 2023	Mentors:	Elisabeth Bullock, Jade Buckwalter, Anushree Gupta, Anne Gvozdjak, Clarise Han, Shauna Kwag, Padmasini Krishnan, Kate Pearce, AnaMaria Perez, Laura Pierson, Swathi Senthil, Jane Wang, Jing Wang, Rebecca Whitman, Muskan Yadav, Valerie Yuen

Some members began creating self-referential multiple-choice tests. We hope that some of these will appear in upcoming issue of the Bulletin soon, so stay tuned.

Other members created their own way of generating random triangles and are now computing the probability that the method produces a right triangle.

Session 32 - Meet 9 April 6, 2023	Mentors:	Elisabeth Bullock, Jade Buckwalter, Anne Gvozdjak, Shauna Kwag, Padmasini Krishnan, Hanna Mularczyk, AnaMaria Perez, Laura Pierson, Vievie Romanelli, Jane Wang, Rebecca Whitman, Muskan Yadav
	Visitor:	Karia Dibert, University of Chicago

Former Girls' Angle mentor Karia Dibert visited us on November 11, 2021 to tell us about a cosmic background radiation detector she designed, destined for installation at the South Pole. She has since made her first visit to the South Pole to begin on site preparations for installation, and she returned to Girls' Angle to tell us about her Antarctic adventure.

Karia's detector will enable a higher resolution image of the cosmic background radiation, more than doubling the number of pixels per region. Later this year, they will install the camera and a test array of pixels, with the full installation projected to take 2-3 years.

The journey to the South Pole itself, was quite an adventure, requiring five flights from Chicago to San Francisco, to Auckland, to Christchurch, to McMurdo Station, and, finally, to the South Pole. There is no runway at the South Pole, so the special airplane, an LC-130, had to be outfitted with special skis in place of wheels.

Karia's colleague Wei Quan showed us some of the cold weather gear they have to wear, including boots with 5" soles and UV light protection, because all the snow reflects a lot of sunlight onto visitors. Even though it's cold and the light shines obliquely, one can get sunburned there.

Karia hopes to collect information on how the universe began, how it has evolved, and about its composition. The detector is also capable of detecting galaxies, asteroids, and star flares, as well as light from any dust around galaxies.

Session 32 - Meet 10 April 13, 2023	Mentors:	Elisabeth Bullock, Anne Gvozdjak, Clarise Han, Shauna Kwag, Padmasini Krishnan, Hanna Mularczyk, Kate Pearce, AnaMaria Perez, Laura Pierson, Swathi Senthil, Jane Wang, Jing Wang, Rebecca Whitman, Muskan Yadav

Some members explored differential equations and how to find approximate solutions to them using Euler's method. Specifically, equations that describe the motion of a pendulum and population growth were examined in detail. Did you know that every day, over 350,000 babies are born throughout the world?

Session 33 - Meet 11 April 27, 2023	Mentors:	Elisabeth Bullock, Jade Buckwalter, Anne Gvozdjak, Shauna Kwag, Hanna Mularczyk, Kate Pearce, AnaMaria Perez, Laura Pierson, Swathi Senthil, Jane Wang, Jing Wang, Muskan Yadav

Given a circle, what are all the circles that intersect it in two points and at right angles? (Two circles intersect at right angles if, at each point of intersection, the tangent lines to each circle through the point of intersection are perpendicular to each other.) If the equation of the given circle is $x^2 + y^2 = 1$, can you give the equations of all the circles that intersect it in this way?

Calendar

Session 31: (all dates in 2022)

September	8	Start of the thirty-first session!
	15	
	22	
	29	
October	6	
	13	
	20	
	27	
November	3	
	10	
	17	
	24	Thanksgiving - No meet
December	1	

Session 32: (all dates in 2023)

January	26	Start of the thirty-second session!
February	2	
	9	
	16	
	23	No meet
March	2	
	9	
	16	
	23	
	30	No meet
April	6	Karia Dibert, University of Chicago
	13	
	20	No meet
	27	
May	4	

Girls' Angle has run over 150 Math Collaborations. Math Collaborations are fun, fully collaborative, math events that can be adapted to a variety of group sizes and skill levels. We now have versions where all can participate remotely. We have now run four such "all-virtual" Math Collaboration. If interested, contact us at girlsangle@gmail.com. For more information and testimonials, please visit www.girlsangle.org/page/math_collaborations.html.

Girls' Angle can offer custom math classes for small groups on a wide range of topics. For inquires, email: girlsangle@gmail.com.

MK, beloved bald eagle of Arlington, passed away on February 28, 2023 from secondary rodenticide poisoning.

Photo courtesy of Margaret Lewis.

R.I.P. MK (2016-2023).

Girls' Angle: A Math Club for Girls

Membership Application

Note: If you plan to attend the club, you only need to fill out the Club Enrollment Form because all the information here is also on that form.

Applicant's Name: (last) ______________________________ (first) ______________________________

Parents/Guardians: ____

Address (the Bulletin will be sent to this address):

Email:

Home Phone: ____________________________ Cell Phone: ____________________________

Personal Statement (optional, but strongly encouraged!): Please tell us about your relationship to mathematics. If you don't like math, what don't you like? If you love math, what do you love? What would you like to get out of a Girls' Angle Membership?

____-

The $50 rate is for US postal addresses only. **For international rates, contact us before applying.**

Please check all that apply:

- □ Enclosed is a check for $50 for a 1-year Girls' Angle Membership.
- □ I am making a tax-free donation.

Please make check payable to: **Girls' Angle**. Mail to: Girls' Angle, P.O. Box 410038, Cambridge, MA 02141-0038. Please notify us of your application by sending email to girlsangle@gmail.com.

A Math Club for Girls

Girls' Angle
Club Enrollment

Gain confidence in math! Discover how interesting and exciting math can be! Make new friends!

The club is where our in-person mentoring takes place. At the club, girls work directly with our mentors and members of our Support Network. To join, please fill out and return the Club Enrollment form. Girls' Angle Members receive a significant discount on club attendance fees.

Who are the Girls' Angle mentors? Our mentors possess a deep understanding of mathematics and enjoy explaining math to others. The mentors get to know each member as an individual and design custom tailored projects and activities designed to help the member improve at mathematics and develop her thinking abilities. Because we believe learning follows naturally when there is motivation, our mentors work hard to motivate. In order for members to see math as a living, creative subject, at least one mentor is present at every meet who has proven and published original theorems.

What is the Girls' Angle Support Network? The Support Network consists of professional women who use math in their work and are eager to show the members how and for what they use math. Each member of the Support Network serves as a role model for the members. Together, they demonstrate that many women today use math to make interesting and important contributions to society.

What is Community Outreach? Girls' Angle accepts commissions to solve math problems from members of the community. Our members solve them. We believe that when our members' efforts are actually used in real life, the motivation to learn math increases.

Who can join? Ultimately, we hope to open membership to all women. Currently, we are open primarily to girls in grades 5-12. We welcome *all girls* (in grades 5-12) regardless of perceived mathematical ability. There is no entrance test. Whether you love math or suffer from math anxiety, math is worth studying.

How do I enroll? You can enroll by filling out and returning the Club Enrollment form.

How do I pay? The cost is \$20/meet for members and \$30/meet for nonmembers. Members get an additional 10% discount if they pay in advance for all 12 meets in a session. Girls are welcome to join at any time. The program is individually focused, so the concept of "catching up with the group" doesn't apply.

Where is Girls' Angle located? Girls' Angle is based in Cambridge, Massachusetts. For security reasons, only members and their parents/guardian will be given the exact location of the club and its phone number.

When are the club hours? Girls' Angle meets Thursdays from 3:45 to 5:45. For calendar details, please visit our website at www.girlsangle.org/page/calendar.html or send us email.

Can you describe what the activities at the club will be like? Girls' Angle activities are tailored to each girl's specific needs. We assess where each girl is mathematically and then design and fashion strategies that will help her develop her mathematical abilities. Everybody learns math differently and what works best for one individual may not work for another. At Girls' Angle, we are very sensitive to individual differences. If you would like to understand this process in more detail, please email us!

Are donations to Girls' Angle tax deductible? Yes, Girls' Angle is a 501(c)(3). As a nonprofit, we rely on public support. Join us in the effort to improve math education! Please make your donation out to **Girls' Angle** and send to Girls' Angle, P.O. Box 410038, Cambridge, MA 02141-0038.

Who is the Girls' Angle director? Ken Fan is the director and founder of Girls' Angle. He has a Ph.D. in mathematics from MIT and was a Benjamin Peirce assistant professor of mathematics at Harvard, a member at the Institute for Advanced Study, and a National Science Foundation postdoctoral fellow. In addition, he has designed and taught math enrichment classes at Boston's Museum of Science, worked in the mathematics educational publishing industry, and taught at HCSSiM. Ken has volunteered for Science Club for Girls and worked with girls to build large modular origami projects that were displayed at Boston Children's Museum.

Who advises the director to ensure that Girls' Angle realizes its goal of helping girls develop their mathematical interests and abilities? Girls' Angle has a stellar Board of Advisors. They are:

At Girls' Angle, mentors will be selected for their depth of understanding of mathematics as well as their desire to help others learn math. But does it really matter that girls be instructed by people with such a high-level understanding of mathematics? We believe YES, absolutely! One goal of Girls' Angle is to empower girls to be able to tackle *any* field regardless of the level of mathematics required, including fields that involve original research. Over the centuries, the mathematical universe has grown enormously. Without guidance from people who understand a lot of math, the risk is that a student will acquire a very shallow and limited view of mathematics and the importance of various topics will be improperly appreciated. Also, people who have proven original theorems understand what it is like to work on questions for which there is no known answer and for which there might not even be an answer. Much of school mathematics (all the way through college) revolves around math questions with known answers, and most teachers have structured their teaching, whether consciously or not, with the knowledge of the answer in mind. At Girls' Angle, girls will learn strategies and techniques that apply even when no answer is known. In this way, we hope to help girls become solvers of the yet unsolved.

Also, math should not be perceived as the stuff that is done in math class. Instead, math lives and thrives today and can be found all around us. Girls' Angle mentors can show girls how math is relevant to their daily lives and how this math can lead to abstract structures of enormous interest and beauty.

Girls' Angle: Club Enrollment Form

Applicant's Name: (last) ______________________________ (first) ______________________________

Parents/Guardians: ____

Address: ____ Zip Code: _________

Home Phone: _________________ Cell Phone: _________________ Email: ______________________

Please fill out the information in this box.

Emergency contact name and number: ____

Pick Up Info: For safety reasons, only the following people will be allowed to pick up your daughter. Names:

____

Medical Information: Are there any medical issues or conditions, such as allergies, that you'd like us to know about?

____

Photography Release: Occasionally, photos and videos are taken to document and publicize our program in all media forms. We will not print or use your daughter's name in any way. Do we have permission to use your daughter's image for these purposes? **Yes** **No**

Eligibility: Girls roughly in grades 5-12 are welcome. Although we will work hard to include every girl and to communicate with you any issues that may arise, Girls' Angle reserves the discretion to dismiss any girl whose actions are disruptive to club activities.

Personal Statement (optional, but strongly encouraged!): We encourage the participant to fill out the optional personal statement on the next page.

Permission: I give my daughter permission to participate in Girls' Angle. I have read and understand everything on this registration form and the attached information sheets.

____ Date: ________________________

(Parent/Guardian Signature)

Participant Signature: ____

Members: Please choose one.	Nonmembers: Please choose one.
☐ Enclosed is $216 for one session (12 meets)	☐ I will pay on a per meet basis at $30/meet.
☐ I will pay on a per meet basis at $20/meet.	☐ I'm including $50 to become a member, and I have selected an item from the left.

☐ I am making a tax-free donation.

Please make check payable to: **Girls' Angle**. Mail to: Girls' Angle, P.O. Box 410038, Cambridge, MA 02141-0038. Please notify us of your application by sending email to girlsangle@gmail.com. Also, please sign and return the Liability Waiver or bring it with you to the first meet.

Personal Statement (optional, but strongly encouraged!): This is for the club participant only. How would you describe your relationship to mathematics? What would you like to get out of your Girls' Angle club experience? If you don't like math, please tell us why. If you love math, please tell us what you love about it. If you need more space, please attach another sheet.

Girls' Angle: A Math Club for Girls
Liability Waiver

I, the undersigned parent or guardian of the following minor(s)

____,

do hereby consent to my child(ren)'s participation in Girls' Angle and do forever and irrevocably release Girls' Angle and its directors, officers, employees, agents, and volunteers (collectively the "Releasees") from any and all liability, and waive any and all claims, for injury, loss or damage, including attorney's fees, in any way connected with or arising out of my child(ren)'s participation in Girls' Angle, whether or not caused by my child(ren)'s negligence or by any act or omission of Girls' Angle or any of the Releasees. I forever release, acquit, discharge and covenant to hold harmless the Releasees from any and all causes of action and claims on account of, or in any way growing out of, directly or indirectly, my minor child(ren)'s participation in Girls' Angle, including all foreseeable and unforeseeable personal injuries or property damage, further including all claims or rights of action for damages which my minor child(ren) may acquire, either before or after he or she has reached his or her majority, resulting from or connected with his or her participation in Girls' Angle. I agree to indemnify and to hold harmless the Releasees from all claims (in other words, to reimburse the Releasees and to be responsible) for liability, injury, loss, damage or expense, including attorneys' fees (including the cost of defending any claim my child might make, or that might be made on my child(ren)'s behalf, that is released or waived by this paragraph), in any way connected with or arising out of my child(ren)'s participation in the Program.

Signature of applicant/parent: ____Date: ______________

Print name of applicant/parent: ____

Print name(s) of child(ren) in program: ___________________________________