To Foster and Nurture Girls’ Interest in Mathematics

An Interview with Sarah Spence Adams, Part 1
The Needell in the Haystack:
Natural Language Processing, Part 3:
From Language Models to AI Chatbots
Generating Functions for Partitions

Summer Fun Problem Sets:
Best Ice Cream Cone
Scissors Congruence
Arts and Graphs
Notes from the Club
From the Founder
An amazing characteristic of mathematical mystery is that, in principle, one has everything one needs to resolve it in one’s head. There’s no need to look outside and no equipment is necessary. Math is highly portable. You can enjoy it just about anywhere. - 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 Dafnaur
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
Dela Cheung Hom and
Eugene Shih
David Kelly
Mark and Lisel Macenka
Brian and Darlene Matthews
Toshia McCabe
Mary O’Keefe
Stephen Knight and
Elizabeth Quattrock 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
Rini Thaman
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.

FOUNDER AND PRESIDENT
C. Kenneth Fan

BOARD OF ADVISORS
Connie Chow
Yaim Cooper
Julia Elisenda Grigsby
Kay Kirkpatrick
Grace Lyo
Lauren McGough
Mia Minnes
Bjorn Poonen
Beth O’Sullivan
Elissa Ozanne
Katherine Paur
Liz Simon
Gigliola Staffilani
Bianca Viray
Karen Willcox
Lauren Williams

On the cover: Ice Cream Cones Galore by C. Kenneth Fan. See Clarise Han’s Summer Fun problem set Best Ice Cream Cone on page 21.
An Interview with Sarah Spence Adams, Part 1

Sarah Spence Adams is Professor of Mathematics and Electrical & Computer Engineering at Olin College. She earned a Bachelor of Science from the University of Richmond and a PhD in Mathematics from Cornell University under the supervision of Stephen B. Wicker.

Ken: You’re an expert on secret codes. Let’s start with a basic question: What do secret codes have to do with math? How does math enter the picture?

Sarah: Secret codes have been around for millennia, and math has been lurking beneath the surface in the majority of known examples. For example, the Caesar cipher, said to be used by Julius Caesar around 100 B.C., shifts each letter forward by three letters so that every A in an original message is replaced by a D, and similarly Bs are replaced by Es, and so on. To shift a Z “forward” by three, it has to “wrap around” to C. Mathematically, we can think about this as performing addition modulo 26, where we consider the letters A to Z as the numbers 0 to 25, respectively, and all addition “wraps around” to keep all numbers between 0 and 25. For example, to shift forward Y by three, we think about Y as 24, and compute $24 + 3 \mod 26$, which means that the normal answer of 27 has to “wrap around” to 1. The number 1 then converts to B, so we see that this Caesar cipher would encrypt Y as B.

LVQ’W WKDW FRRO?

Modern secret codes, or cryptographic ciphers, use more advanced mathematical ideas/structures such as modular exponentiation, discrete logarithms, and elliptic curves. One really cool thing is that modular arithmetic is at the core of these more advanced topics. It is fascinating to me that modular arithmetic is found in both ancient, easy-to-break ciphers and modern, highly secure ciphers.

Ken: What exactly is coding theory? What are the “big problems” in coding theory?

Sarah: Whereas cryptography involves secret codes that are used to keep information safe from intruders or eavesdroppers, coding theory involves error control codes that are used to keep information safe from the inevitable errors that occur during the transmission of a message or during the storage and later retrieval of data. For example, a message traveling to Earth from a deep space satellite might experience errors due to a solar flare or other atmospheric disturbances, while data being retrieved from a DVD might experience errors due to fingerprints or scratches on the DVD.

Error control codes insert redundancy into the original message so that the intended message can be recovered even if parts of it are corrupted by errors. A silly example would be to consider the telephone game, where one friend begins by whispering a sentence into another friend’s ear. This sentence is passed around a room via whispers, and usually, the last person says a sentence out loud that is wildly different from the original sentence. If we changed this game and allowed each person to repeat the sentence five times, we would expect that this repetition would help reduce the number of errors that get introduced and propagated around the room. Error control
codes do something similar, however they can’t simply repeat each message five times as that would slow down our communications too much.

The big problem in coding theory is to figure out how to introduce redundancy in clever ways, tailored to the needs of a particular application. There are many factors to consider when creating or choosing error control codes for particular channels, including the expected frequency of errors, the expected distribution of errors (spread out or clumped together?), how much power or bandwidth is available, how fast we need the transmissions to be, and whether the communication channel is “one-way” or if we can ask for a repeat transmission if something goes wrong. Coding theory is a big field that requires cooperation among mathematicians, engineers, and telecommunications experts.

Ken: Could you please explain one of the more sophisticated mathematical concepts that you use in your work and how it is used?

Sarah: Some of the most powerful and beautiful error control codes are built using ideals of polynomial rings over Galois fields. I learned about these structures as an undergraduate, and then I concentrated on codes built using these structures during graduate school. I love applying these beautiful mathematical structures to real-world problems.

To understand ideals of polynomial rings over Galois fields,¹ we need to first understand the concept of a field. In this interview, I will keep things as informal as possible, so you are encouraged to look at an abstract algebra book to see the formal details. A field is a special collection of elements, along with two operations, that behave in nice ways. The real numbers form a field under addition and multiplication due to the nice properties about how real numbers interact with each other. For example, real numbers follow the associativity, commutativity, and distributivity properties for both addition and multiplication. Also, there is a special element, namely 0, that “does nothing” under addition. For any real number $x$, we know $x + 0 = 0 + x = x$. Similarly, there is a special element, namely 1, that “does nothing” under multiplication. For any real number $x$, we have $x(1) = 1(x) = x$. These special numbers are called additive and multiplicative identities, respectively.

Finally, every real number $x$ has an “additive inverse,” meaning a number that can be added to $x$ to obtain the additive identity (0). For example, 5 has an additive inverse of -5 because $5 + (-5) = 0$. Similarly, every nonzero real number $x$ has a “multiplicative inverse,” meaning a number that can be multiplied by $x$ to get back to the multiplicative identity (1). For example, 12 has a multiplicative inverse of 1/12 because $(12)(1/12) = 1$. In summary, the real numbers under the operations of addition and multiplication form a field because those operations satisfy the associative, commutative, and distributive properties, they have additive and multiplicative identities, and all real numbers have an additive inverse, while all nonzero real numbers also have a multiplicative inverse.

Sometimes it helps to understand a definition by considering a counterexample. The integers under addition and multiplication do not form a field. Although they satisfy most of the requirements for a field, multiplicative inverses do not exist for all integers. For example, consider the

---

¹ Galois fields are finite fields. For a Learn by Doing on finite fields, see page 23-26 of Volume 9, Number 1 of the Girls’ Angle Bulletin. Galois fields are named after Évariste Galois, a French mathematician who lived from 1811 to 1832.
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. Sarah Spence Adams 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
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](http://www.ams.org/bookstore-getitem/item=MBK-84-90) and use the code “GIRLS” at checkout.
America’s Greatest Math Game: Who Wants to Be a Mathematician.

(advertisement)
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
You have probably heard about or encountered the wonders of ChatGPT, an artificial intelligence chatbot developed by OpenAI. ChatGPT has been shown to do some pretty incredible things: it can help you outline an essay, explain how to make a peanut butter sandwich from the perspective of a pirate, or even write a story about Natural Language Processing (NLP), a dragon, and a Princess. But how does it work? Is there a human who sits behind the scenes and answers prompts, Wizard-of-Oz style? That’s not likely, especially with the breadth of topics and speed at which ChatGPT responds to prompts. In this installment of “Needell in the Haystack” (“NITH”), we will see how language models can be used to generate text. We’ll review a simple language model discussed in the last installment of “NITH” and then generalize to the n-gram model. Finally, we will discuss how many modern chatbots such as ChatGPT work by contrasting simpler models with the Transformer Model.

A Simple Language Model

Before diving in, let’s review some fundamental concepts of language models we’ve discussed previously. Language models are computational models used to generate human text. We first recap a simple model that takes a text data set, splits it into pairs of words known as bi-grams, and trains a transition matrix. These steps are summarized and reviewed in Figure 1 and Figure 2. This model is known as the bi-gram model because of the use of word pairs.

Figure 1. The text dataset is split into word pairs, or bi-grams, to train the transition matrix.

<table>
<thead>
<tr>
<th>START</th>
<th>I</th>
<th>my</th>
<th>math</th>
<th>think</th>
<th>mango</th>
<th>is</th>
<th>very</th>
<th>squishy</th>
<th>fun</th>
<th>problems</th>
<th>and</th>
<th>has</th>
<th>challenging</th>
<th>END</th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
<td>1/2</td>
<td>1/4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>I</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>my</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>math</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>think</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mango</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>is</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>very</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>squishy</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>fun</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>problems</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1/3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2/3</td>
</tr>
<tr>
<td>and</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>has</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>challenging</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 2. The elements of the transition matrix are determined by counting the number of bi-grams that start with word I and end with word J, and dividing by the number of word pairs that start with word I. For example, the entry in row I = “math” and J = “is” is 2/3 because three bi-grams start with the word “math” (highlighted in red in Figure 1) and two of them end with the word “is.”

1 This content is supported in part by a grant from MathWorks. Anna Ma is an Assistant Professor at the University of California Irvine.
3 The ‘GPT’ in ChatGPT stands for Generative Pre-trained Transformer. ChatGPT is a transformer model that generates text based on data it was previously trained with.
The transition matrix contains conditional probabilities between words. For instance, 
\( P(\text{is} \mid \text{math}) = \frac{2}{3} \), highlighted in red in Figure 2, represents the probability that “is” appears next conditional on “math” appearing before it. The transition matrix can be used to assign probabilities to entire sequences of words by multiplying the conditional probabilities appearing in a sequence. For example, the transition matrix in Figure 2 assigns a probability of \( \frac{4}{54} \) to the sentence “I think math is fun,” because

\[
P(\text{START I think math is fun END})
= P(\text{START})P(\text{I} \mid \text{START})P(\text{think} \mid \text{I})P(\text{math} \mid \text{think})P(\text{is} \mid \text{math})P(\text{fun} \mid \text{is})P(\text{END} \mid \text{fun})
= 1 \times \frac{1}{2} \times 1 \times 1 \times \frac{2}{3} \times 1 \times \frac{1}{3} = \frac{4}{54},
\]

where we use Bayes’ Theorem, which says \( P(\text{EF}) = P(\text{F})P(\text{E} \mid \text{F}) \) where \( \text{E} \) and \( \text{F} \) are words.

Unfortunately, this transition matrix assigns an even larger probability to the sentence “I think math is squishy,” since

\[
P(\text{START I think math is squishy END})
= P(\text{START})P(\text{I} \mid \text{START}) \cdot P(\text{think} \mid \text{I}) \cdot P(\text{math} \mid \text{think}) \cdot P(\text{is} \mid \text{math}) \cdot P(\text{squishy} \mid \text{is}) \cdot P(\text{END} \mid \text{squishy})
= 1 \times \frac{1}{2} \times 1 \times 1 \times \frac{2}{3} \times 1 \times 1 = \frac{2}{18}.
\]

This highlights a significant issue with our current model.

**Drawbacks of the Current Model**

In any data science application and with any machine learning model, it’s important to understand what our method can and cannot do. While this bi-gram model can be used to complete sentences, the (minimal) text data set it was trained on only mentions mangoes and math. Thus, it cannot complete sentences about other topics, nor sentences such as

“*There is no place like home. There is no place like ___*”

If we want a model that can generate more varied sentences, we will need more training data. However, even with more data, this model is still limited because the probability of the next word is only dependent on the preceding word. Here’s an example to highlight this limitation:

“*My favorite subject in school is math and ___*”

The sentence is about math, so one might expect the current model to be able to complete this sentence (assuming additional data about math is added to the data set). Still, the model doesn’t understand context. The context of the sentence is “favorite subjects”; however, to complete the sentence, the bi-gram model only pays attention to the word “and” when determining the next word. In the next section, we will discuss the generalization of the bi-gram model: the **n-gram model** is a model that moves us a step closer to contextual awareness.
The N-Gram Model

The \( n \)-gram model is a generalization of the bi-gram model in which the previous \( n - 1 \) words are used to determine the next word, instead of only using the previous word to determine the next word. This is accomplished by conditioning not just on the preceding word, but on the \( n - 1 \) preceding words. We call this a generalization because the bi-gram model is a special case of the \( n \)-gram model when \( n = 2 \). In Figure 1, we broke up our text data set into bi-grams. For an \( n \)-gram model, we would break up the text into \( n \)-grams, as shown in Figure 3.

The transition matrix for a 4-gram still contains conditional probabilities, except here, we conditioned on the preceding three words. The transition matrix for the 4-gram model is given in Figure 4. For example, the conditional probability that “is” follows “I think math,” \( P(\text{is} | \text{I think math}) \), is 1 since there are two 4-grams that start with “I think math” and both end with “is” (highlighted in red in Figure 3). We can also compute

\[
P(\text{START I think math is fun END}) = P(\text{START I think}) \cdot P(\text{math | START I think}) \cdot P(\text{is | I think math}) \cdot P(\text{fun | think math is}) \cdot P(\text{END | math is fun}) = \frac{1}{2} \times 1 \times 1 \times \frac{1}{2} \times 1 = \frac{1}{4},
\]

where \( P(\text{START I think}) = \frac{1}{2} \) because two out of the four sentences in the data set start with “START I think.”

The conditional probability that “squishy” follows “think math is” is 0 since there are no 4-grams that start with “think math is” and end with “squishy.” In particular, “squishy” does not appear in the context of “math” in our data set, and the 4-gram model allows us to capture this. We can also compute

\[
P(\text{START I think math is squishy END}) = P(\text{START I think}) \cdot P(\text{math | START I think}) \cdot P(\text{is | I think math}) \cdot P(\text{squishy | think math is}) \cdot P(\text{END | math is squishy}) = \frac{1}{2} \times 1 \times 1 \times 0 \times 0 = 0,
\]

where \( P(\text{END | math is squishy}) = 0 \) since “math is squishy” doesn’t precede any word in our original data set (and thus does not have a row in the transition matrix).

What do you notice when comparing the matrices in Figure 2 and Figure 4? For one thing, the 4-gram transition matrix has fewer columns. The 4-gram transition matrix also has less variation in its entries and is more sparse (more zero entries); its values are only 0, 1, or \( \frac{1}{2} \). Combined, this means that the 4-gram model generates fewer unique sentences in comparison to the bi-gram model. In fact, you can verify that the 4-gram model only assigns nonzero probabilities to four sentences, all with probability \( \frac{1}{4} \):
1. “I think math is very fun”  
2. “I think math is fun”  
3. “My mango is squishy”  
4. “Math has challenging and fun problems”  

Do these sentences seem familiar? Compare them with the sentences in Figure 3. While the 4-gram model may capture contextual information, more data is needed to generate unique sentences that aren’t from the training data set. This highlights the trade-off of using the bi-gram model versus the 4-gram model.

<table>
<thead>
<tr>
<th></th>
<th>math</th>
<th>is</th>
<th>very</th>
<th>fun</th>
<th>squishy</th>
<th>challenging</th>
<th>and</th>
<th>problems</th>
<th>STOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>START I think</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>I think math</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Think math is</td>
<td>0</td>
<td>0</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Math is very</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Is very fun</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Math is fun</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>START my mango</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>My mango is</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mango is squishy</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>START math has</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Math has challenging</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Has challenging and</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Challenging and fun</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>And fun problems</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 4. The transition matrix for the 4-gram model.

**The Transformer Model**

ChatGPT uses a more sophisticated model: the transformer model. A **transformer model** is a series of functions that can output a sequence of words. Unlike the methods we’ve already discussed, the manner in which probabilities are assigned is not as straightforward, but this won’t stop us from getting a general idea of how the transformer model works. Figure 5 provides a visualization of the $n$-gram model (top) and a simplified transformer model (bottom) so that we can compare the two.

Let’s first put the $n$-gram model with $n = 2$ into the context of Figure 5. The input for the bi-gram model is the preceding word, and the output is the next chosen word. For example, if you want to complete the sentence, “I think math is,” then “is” is the input to the model. Next, the transition matrix takes the word “is” and checks which words will likely appear after “is.” One of these words, for example, “fun,” is then selected and returned as the output.

The main similarities between $n$-gram and transformer models are that they have an input, an output, and a step that determines the probabilities of the next word appearing. In the bi-gram model the input is one word and the output is one word, whereas in a transformer model, the input can be a sequence of words and the output another sequence of words. The transition matrix in the bi-gram model tells us the conditional probabilities between pairs of words. The transformer model replaces the transition matrix step with a **neural network**. For simplicity, we
will consider neural networks as a black box function that takes some input and returns a probability. Instead of discussing the neural network portion, let’s shift our attention to another key difference between the two approaches.

![Diagram of n-gram model and transformer model](image)

Figure 5. (Top) n-gram model and (bottom) transformer model.

The **attention** step of the transformer model is what allows it to draw relationships of words in the context of a sentence. For a given sequence of words, the attention mechanism places different weights on each word; words that are important to the context of the sentence get larger weights than words that are not important to the context of the sentence. For example, consider the sentence

“My favorite subjects in high school are math and …. .”

The attention step tells the model to use the fact that the next word should be related to “subjects” in “high school.” This is just one of many reasons transformers can perform so well. Most large-scale transformer models even use multiple attention functions to capture different levels of abstraction of a sequence of text.

**Concluding Remarks**

In today’s fast-paced world, language models are of great potential, but understanding what’s happening under the hood and the limitations of technologies, such as ChatGPT, is essential. Language models consist of probabilities based on a large corpus of training data. Thus – at least when using language models – any chatbot will be limited to the quality and/or quantity of data used to train it. In addition to the quality and quantity of data, how that data is being used in the model (e.g., bi-gram versus n-gram) heavily impacts how we should interpret and how much we should rely on the output of the model. The key to comprehending these limitations is understanding the foundational mathematics and assumptions beneath the models. I hope the “NITH” NLP series has given you an extraordinary glimpse into the world of chatbots!
Generating Functions for Partitions
by Robert Donley
edited by Amanda Galtman

In the previous installment, we developed counting methods for compositions and introduced partitions. These concepts differ only in a notion of ordering, so we might not expect significant differences in theory. On the contrary, partitions are a topic of intense perennial interest and occupy a central role in several fields within mathematics and physics.

We begin with the generating function approach to partitions, with a focus on partial sums, partial fractions, and the binomial series. With these techniques, we find that partitions are counted by functions called quasipolynomials. We’ll explain what a quasipolynomial is later, when the context will make the notion clearer.

**Definition.** A partition of a positive integer \( k \) into \( m \) parts is a sum of entries \( x_1 + \ldots + x_m = k \) where each \( x_i \) is a positive integer and \( x_i \geq x_i + 1 \). (The \( x_i \) are the parts of the partition.)

When all parts are single-digit numbers, we simply list the parts as a string of digits.

We define the function \( p(k) \) to be the number of partitions of \( k \) of any length, while \( p_{[n]}(k) \) denotes the number of partitions of \( k \) with parts in 1, ..., \( n \). As a useful convention, we define \( p(0) \) and \( p_{[n]}(0) \) to be 1.

**Example:** As strings of digits, the partitions of 5 are given by

\[
5 \quad 41 \quad 32 \quad 311 \quad 221 \quad 2111 \quad 11111
\]

Thus, for example, \( p(5) = 7 \) and \( p_{[3]}(5) = 5 \).

We counted compositions using generating functions with several variables. Similar methods apply here, but since no ordering is required to identify a partition, we simply count using exponent rules with a single variable. For example, the generating function for \( p_{[1]}(k) \) is

\[
p_{[1]}(0) + p_{[1]}(1)t + p_{[1]}(2)t^2 + \ldots = 1 + t + t^2 + \ldots = \frac{1}{1-t}.
\]

The second equality is the geometric series formula. Since only strings of 1s are allowed, the exponent merely records the number of 1s.

For the generating function of \( p_{[2]}(k) \), we multiply the generating function for \( p_{[1]}(k) \) by the factor that generates all multiples of 2 to get

\[
p_{[2]}(0) + p_{[2]}(1)t + \ldots = (1 + t + t^2 + \ldots)(1 + t^2 + t^4 + \ldots) = \frac{1}{1-t} \cdot \frac{1}{1-t^2}.
\]

---

1 This content is supported in part by a grant from MathWorks.
In general, the generating function for $p_{[n]}(k)$ is given by

$$p_{[n]}(0) + p_{[n]}(1)t + p_{[n]}(2)t^2 + \ldots = \frac{1}{1-t} \cdot \frac{1}{1-t^2} \cdot \frac{1}{1-t^n},$$

and the generating function for all partitions is given by the infinite product

$$p(0) + p(1)t + p(2)t^2 + \ldots = \frac{1}{1-t} \cdot \frac{1}{1-t^2} \cdot \frac{1}{1-t^3} \cdot \ldots.$$

Each additional factor includes the corresponding part with all its multiplicities. To find the coefficient of $t^k$ in the general formula, we require only the first $k$ terms of the product, and of these, only the terms up to $t^k$ contribute.

**Example:** The levels of the Hasse diagram below give $p_{[2]}(k)$ for $0 \leq k \leq 5$. We confirm by multiplying

$$(1 + t + t^2 + t^3 + t^4 + t^5)(1 + t^2 + t^4) = 1 + t + 2t^2 + 2t^3 + 3t^4 + 3t^5 + \ldots.$$

**Exercise:** Verify that each partition corresponds to a unique set of exponents in the product above. For instance, 2111 arises from the exponents in the product $t^3 \cdot t^2 = t^{3+1} \cdot t^{1+2}$ as three 1s and one 2.

**Exercise:** Expand $\frac{1}{1-t} \cdot \frac{1}{1-t^2} = \frac{1}{(1-t)^2} \cdot \frac{1}{1+t}$ as a partial fraction, apply the geometric series to each term, and combine like terms. Then, extract the coefficient of $t^k$ to find the complete formula for $p_{[2]}(k)$.

It is convenient to write the formula based on remainders. Express $k$ as $2s + r$, where $s$ and $r$ are integers and $r$ is 0 or 1, depending on whether $k$ is even or odd, respectively. In this notation, $p_{[2]}(2s) = s + 1$ and $p_{[2]}(2s + 1) = s + 1$.

Here’s another way to calculate $p_{[2]}(k)$. If we multiply the generating function for a sequence $\{a_k\}$ by $1/(1-t)$, then the result gives the generating function for the sequence of partial sums $s_k = a_0 + \ldots + a_k$.

**Exercise:** Obtain $p_{[2]}(k)$ using partial sums. That is, expand the geometric series for $1/(1+t)$, take the partial sums, and then take the partials sums again.

**Exercise:** On the other hand, note the special case of the binomial series

$$1 + 2t + 3t^2 + 4t^3 + \ldots = \frac{1}{(1-t)^2}.$$

How do we adjust the partial sums if we multiply this series by $1 - t + t^2 - t^3 + \ldots$?
To append factors for larger parts to the generating function, if we multiply by

$$1 + t + t^2 + \ldots = \frac{1}{1-t},$$

then we obtain aerated partial sums, or sums with uniform gaps.

**Example:** We multiply the two factors $1/(1-t)$ and $1/(1-t^2)$ for $p_{12}(k)$ using an aerated partial sum. If we multiply the generating function for $\{a_k\}$ by $1/(1-t^2)$, then we obtain the generating function with coefficients $s_k = a_k + a_{k-2} + a_{k-4} + \ldots$.

For instance, the following table shows a calculation of $p_{12}(8)$:

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

**Exercise:** Fix $n \geq 3$. Find the formula for the number of partitions with parts 1 and $n$. Explain both by using partial sums and by listing the partitions directly.

This process works in general to calculate values for other $p_{(n)}(k)$; we account for each new factor by applying an aerated partial sum. To obtain $p_{(n)}(k)$, we sum the value directly above with entries to its left, leaving gaps of size $n-1$. For instance, we can obtain small values for $p_{[3]}(k)$ and $p_{[4]}(k)$ using the following tables:

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>9</td>
<td>11</td>
<td>15</td>
</tr>
</tbody>
</table>
Exercise: Confirm all counts in the last table by finding the corresponding partitions. Then repeat using aerated partial sums. Find \( p_{[3]}(k) \) up to \( k = 23 \). A later exercise will use these values.

Exercise: The values on the diagonal starting in column 1 stabilize to values of \( p(k) \). Explain this fact using the recursive rule.

A three-term recursive formula is given by \( p_{[n]}(k) = p_{[n-1]}(k) + p_{[n]}(k-n) \).

Exercise: Prove the three-term recurrence, either by considering parts or by manipulating generating functions. Then explain the mechanics of the recurrence in the tables above. Use the recurrence to verify the above tables.

Exercise: Iterate the three-term recurrence by replacing the second term until it naturally stops, and interpret as a sum along a diagonal. What happens if we iterate by replacing the third term?

We now have the basic techniques for our main task: to find the general formula for \( p_{[3]}(k) \).

First, let’s rewrite the generating function for \( p_{[3]}(k) \) as

\[
\frac{1}{1-t} \cdot \frac{1}{1-t^2} \cdot \frac{1}{1-t^3} = \frac{1}{(1-t)^3(1+t)(1+t^2)} = \frac{1-t+t^2}{(1-t)^2(1-t^3)}.
\]

Exercise: Verify the second equality directly or by factoring \( 1 - t^6 \). To obtain the quadratic factors, factor \( 1 + t^3 \) and \( 1 - t^3 \).

To proceed as before, we take two partial sum sequences of the numerator and then apply aerated partial sums with gap length 5.

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – ( t ) + ( t^2 )</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>partial sums of row above</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>partial sums of row above</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>

Except when \( k = 0 \), the values in the third line equal \( k \), and an argument similar to the above gives the recurrence

\[ p_{[3]}(k) = k + p_{[3]}(k-6). \]

Exercise: Verify the recurrence, and describe the implementation in the table.

Here’s where remainders simplify our work. For example, consider the values for \( k = 6s + r \), where \( s \) and \( r \) are integers with \( 0 \leq r < 6 \). Then the recurrence becomes
Now define \( p_{[3], r}(s) = p_{[3]}(6s + r) \), so the recurrence becomes

\[
p_{[3], r}(s) = (6s + r) + p_{[3], r}(s - 1).
\]

Let’s see what pattern results if we iterate the recurrence:

\[
p_{[3], r}(0) = r,
p_{[3], r}(1) = (6 + r) + r
\]

\[
p_{[3], r}(2) = (6 \cdot 2 + r) + (6 + r) + r,
p_{[3], r}(3) = (6 \cdot 3 + r) + (6 \cdot 2 + r) + (6 + r) + r.
\]

In general,

\[
p_{[3], r}(s) = (6s + r) + \ldots + (6 \cdot 2 + r) + (6 + r) + r
\]

\[
= 6(1 + 2 + 3 + \ldots + s) + (s + 1)r
\]

\[
= 3s(s + 1) + (s + 1)r.
\]

If we simplify and rename, we have \( p_{[3]}(6s + r) = 3s^2 + (r + 3)s + r \). For instance, with \( r = 2 \), we have \( p_{[3]}(6s + 2) = 3s^2 + 5s + 2 \). When a function, such as \( p_{[3]}(k) \) is given by different polynomials for different remainder classes (i.e., different values of \( r \) in this case), we call the function a **quasipolynomial**.

**Exercise:** Find the formula for the quadratic polynomial \( p_{[3], 0}(s) \) with \( k = 6s \). Why do we consider this case separately? Verify these six polynomials by evaluating up to \( k = 23 \).

Why not use the original recurrence? This way gives a polynomial count for each remainder class directly. The term \( p_{[3]}(k - 3) \) suggests fewer remainders to consider, but the formula for \( p_{[2]}(k) \) splits into two cases, resulting in the same six cases for \( p_{[3]}(k) \).

**Exercise:** Rederive these polynomials using the original recurrence.

**Exercise:** Find a single formula for \( p_{[3]}(k) \) by expanding the generating function using partial fractions and geometric series. You should at least note how the terms match the partial fraction expansion with unknown constants.

**Answer:**

\[
p_{[3]}(k) = \frac{1}{6} \binom{k + 2}{3} + \frac{1}{4} \binom{k + 1}{1} + \frac{17}{72} + \frac{1}{8} (-1)^k + \frac{2}{9} \cos\left(\frac{2k\pi}{3}\right).
\]

Here, \( \binom{a}{b} \) denotes the binomial coefficient, “\( a \) choose \( b \).”

The roots of \( t^2 + t + 1 \) require complex numbers and contribute a cosine term to the final formula. The sum of the final two terms always has absolute value less than 1/2, so we obtain the same result if we sum the first three terms and round to the nearest integer.

**Exercise:** Verify this formula and approximation rule for \( k \) up to 11. Expand each term to verify the polynomials for each remainder for \( n = 6 \).
The generating function for general $p_{[a]}(k)$ belongs to a special class of rational functions. The roots of the denominator consist only of roots of unity; these roots are solutions to some $t^m - 1 = 0$ and have the form

$$t = e^{2\pi ki/m} = \cos \frac{2\pi k}{m} + i \sin \frac{2\pi k}{m}.$$ 

This property of the denominator guarantees the quasipolynomial behavior of $p_{[a]}(k)$. We highlight some features of quasipolynomials in the following exercise sequence.

**Exercise:** Fix $n \geq 1$. Recall that the $n$th binomial series is given by

$$\frac{1}{(1-t)^n} = 1 + \binom{n}{1}t + \binom{n+1}{2}t^2 + \ldots.$$

Verify that $f(k) \equiv \binom{n+k-1}{k}$ is a polynomial in $k$.

**Exercise:** Fix $n \geq 1$. Define

$$\frac{1}{(1-t^n)^n} = f(0) + f(1)t + f(2)t^2 + \ldots.$$ 

Use the binomial series to explain why $f(mk + r)$ is a polynomial for each remainder $r$ upon division by $m$.

**Exercise:** Consider

$$\frac{1}{(1-t^2)(1+t)^2} = f(0) + f(1)t + f(2)t^2 + \ldots.$$ 

Verify that $f(4s + r)$ is a polynomial in $s$ for $r = 0, 1, 2, 3$. Express the left-hand side in the form of a rational function with denominator as in the previous exercise.

**Exercise:** Derive the formula for $p_{[2]}(k)$ from the generating function $\frac{1+t}{(1-t^2)^2}$. Repeat for $p_{[3]}(k)$.

These final exercises consider some variations on the number of parts.

**Exercise:** Repeat the general analysis for the number of partitions with parts equal to 2 or greater. Find the polynomials and general formula for the case with only parts 2 and 3.

**Exercise:** Find the polynomials and general formula with parts 1, 2, and 4. How about parts 2, 3, and 4? How far can you get with parts 1, 2, 3, and 4?
The best way to learn math is to do math. Here are the 2023 Summer Fun problem sets.

We invite all members and subscribers to send any questions and solutions to us at girlsangle@gmail.com. We’ll give you feedback and might put your solutions in the Bulletin!

In the August issue, we will provide some solutions. You could wait until the August issue to see the answers, but you will learn a lot more if you try to solve these problems on your own.

Some problems are very challenging and could take several weeks to solve, so please don’t approach these problems with the idea that you must solve them all. Our main goal is to give you some interesting things to think about.

If you get stuck, try to formulate a related question that you can see a way to actively explore to get your mind moving and your pencil writing. If you don’t understand a question, email us.

If you’re used to solving problems fast, it can be frustrating to work on problems that take weeks to solve. Try to enjoy the journey and don’t be afraid to follow detours. It’s like hiking up a mountain. Getting to the top rewards one with a spectacular view, but during the journey, there’s a lot to see and experience. So here’s a meta-problem for those of you who feel frustrated when doing these problems: see if you can dissolve that frustration and replace it with a relaxed, optimistic sense of adventure!

This is Summer Fun, not Summer Torture!
Best Ice Cream Cone
by Clarise Han

Coco wants to make the best ice cream cone ever and is choosing between the following yummy ice cream flavors and toppings:

Flavors: vanilla, chocolate, strawberry, mint chocolate, rocky road, cookie dough, matcha.

Toppings: strawberry bits, popping boba, Oreos, gummy bears, marshmallows, chocolate chips, sprinkles, hot fudge, caramel sauce.

1. Coco wants a cone with three stacked scoops of ice cream. How many possibilities are there? (For ice cream connoisseurs, the scoop order in the stack matters!)

2. Coco wants all three scoops to be of different flavors. Now how many possibilities are there?

3. Coco decides she wants five scoops, all different, stacked in her cone. Now how many possibilities are there?

4. Coco loves the combination of mint chocolate and cookie dough and insists on having them. Considering all of the scoops are of different flavors and the maximum number of scoops is five, how many different cones can Coco make that include mint chocolate and cookie dough?

5. If all five scoops are of different flavors, and mint chocolate and cookie dough must be included in the ice cream cone, what is the probability that the mint chocolate and cookie dough scoops are touching each other?

6. Coco is choosing her toppings. How many different three-topping combinations are there? (The three toppings in the mixture must all be different.)

7. Coco wants a five-topping combination that includes at least one of hot fudge or caramel sauce. How many different toppings are possible?

8. How many ice cream cones can Coco make with three different flavors of ice cream and up to five toppings, providing that one of the flavors is matcha and one of the toppings is popping boba?

9. Coco and her friend Didi both make a seven-stack ice cream cone using all seven flavors. Assuming each possible stack is equally likely, what is the probability that no two flavors occupy the same positions in the two stacks?
Scissors Congruence
by Matthew Bates and Jane Wang

A polygon is figure in the plane made up of straight segments that connect to enclose an area. For simplicity, we will restrict our attention to polygons that do not self-intersect. Here are some examples of polygons:

Here is a question about cutting up shapes: Imagine that you have two polygons. When is it possible to cut up the first polygon via finitely many straight cuts and rearrange the pieces to form the second polygon. When two polygons can be rearranged as such, we call them scissors congruent. Here are some examples of polygons that are scissors congruent:

Let’s start by getting a feel for scissors congruence.

1. Show that the following pairs of shapes are scissors congruent by finding a way to cut and rearrange the first shape to the second. Each shape is displayed on a grid so that you can see the scale of the shape.

2. Justify why two shapes that do not have the same area cannot be scissors congruent.

It turns out that different area is the only thing that stops two polygons from being scissors congruent. We will show this amazing fact in the next series of exercises. Note that the exercises are independent from each other, so if you’re stuck on one, try the other ones and come back to it later!
3. Prove that every triangle is scissors congruent to a rectangle.

4. In this exercise, we will prove that every rectangle is scissors congruent to a square.

   (a) Suppose that we have a rectangle of width \( w \) and height \( h \) with \( h \leq w \leq 4h \). Superimpose a square of area \( hw \) onto the rectangle so that their bottom left corners match, and connect the top left corner of the square to the bottom right corner of the rectangle (as shown below). Consider the following proposed decomposition of the rectangle into a square of equal area.

   ![Rectangle and Square Diagram]

   Show that this is indeed a scissors congruence of the rectangle into an equal area square by:

   i. proving that the rectangle pieces and square pieces are congruent, and
   ii. figuring out where the condition \( h \leq w \leq 4h \) was used.

   (b) Show that every rectangle is scissors congruent to a rectangle of width \( w \) and height \( h \) for which \( h \leq w \leq 4h \).

   (c) Use the results of parts 4a and 4b to prove that every rectangle is scissors congruent to a square of the same area.

5. In this exercise, we will show that the union of two squares is scissors congruent to a single square. Consider the following figure:
(a) Prove that this is a scissors congruence of the union of two squares to a single square.

(b) Explain why this scissors congruence gives a proof of the Pythagorean theorem that \( a^2 + b^2 = c^2 \) in any right triangle with side lengths \( a \) and \( b \) and hypotenuse \( c \).

6. Show that scissors congruence is an **equivalence relation**. That is, show that it satisfies the following three properties:

   (a) Every polygon is scissors congruent to itself.

   (b) If polygon \( P \) is scissors congruent to polygon \( Q \), then \( Q \) is scissors congruent to \( P \).

   (c) Let \( P, Q, \) and \( R \) be three polygons. If \( P \) is scissors congruent to \( Q \) and \( Q \) is scissors congruent to \( R \), then \( P \) is scissors congruent to \( R \).

7. A **triangulation** of a polygon is a way of cutting up a polygon into finitely many triangles. Here are two examples of triangulations of a hexagon:

![Hexagon Triangulation](image)

Draw triangulations on each of the three polygons in the introduction.

8. Show that every polygon has a triangulation.

   **[Hint 1: The vertices of the triangles do not all have to be vertices of the original polygon.]**

   **[Hint 2: Can you start by showing that you can decompose every polygon into trapezoids?]**

9. The following exercise is not necessary for our scissors congruence proof, but is an interesting challenge: Show that every polygon has a decomposition into triangles where the vertices of each triangle lie on vertices of the polygon (e.g. the right triangulation of the hexagon in Problem 7 satisfies this property, but the left one does not).

   **[Hint: it may help to induct on the number of sides of the polygon.]**

10. Use the results of the previous exercises to show that any two polygons with the same area are scissors congruent to each other.
Arts and Graphs
by Hanna Mularczyk

Graph Drawing

A (simple) graph is a collection of vertices along with a collection of edges, which are pairs of distinct vertices. Thus, to define a graph $G$ we simply need to list the set of vertices, $V$, and the set of edges, $E$.

Example 1: $V = \{1, 2, 3, 4, 5, 6, 7\}$, 
$E = \{(1, 2), (1, 4), (1, 5), (2, 3), (2, 4), (2, 6), (4, 6), (5, 6), (6, 7)\}$

If you’ve worked with graphs before, you know that this is not how we usually present them. Graphs are meant to be drawn! And as we are merely human, we often draw them on the most readily-available medium: 2-dimensional paper, or for the more pedantic mathematician, a 2-dimensional plane.

How opportune that we have such a piece of paper at this very moment! Let’s draw the above graph. We draw vertices as dots and edges as lines between the two vertices that define them. Here is a drawing of the graph from Example 1.

There is ultimately only one way to list the vertices and edges of a graph (up to relabeling and re-ordering). But there are many ways to draw a graph!

1. Draw the graph from Example 1 in a way that’s different to how I did.

Intentionally or unintentionally, when we draw a graph, we often take aesthetic considerations into account. This may come as a surprise to those who think mathematicians live robot-like existences, but like any other kind of drawing, graph drawing is an art.

2. Before continuing, reflect on what kinds of decisions you made when making your drawing in the previous question and what kinds of qualities you would strive for in a graph drawing. Then, taking into account the subjective nature of the italicized words, draw the same graph as above:

   a. In the nicest way you can
   b. In the least nice way you can
   c. In the most line-like way you can
   d. In the most random way you can
Planar Graphs

If you saw that I drew the graph from above in the following way (at right), you might be less satisfied with it, compared to my initial drawing. Why? Likely because I drew it in such a way where edges intersect in places that aren’t vertices. For example, the edge (1, 5) crosses (2, 4), (2, 3), and (2, 6). We call a drawing of a graph **planar** if no edges cross each other. We call a graph itself **planar** if there exists a planar drawing of it. The example graph we’ve been using so far is planar, as proven by my first drawing, even though there are some drawings of it that have crossings, like my last drawing. Remember we only need one planar drawing for the graph to be planar!

3. For each of the following graphs, decide whether it is planar or not. If a graph is planar, draw a planar drawing of it. If a graph is not, explain why a planar drawing can’t exist.

   a. $V = \{1, 2, 3, 4\}$, $E = \{\text{every pair of vertices}\}$ (the **complete graph** on 4 vertices)
   b. The complete graph on 5 vertices
   c. $V = \{1, 2, ..., 6\}$, $E = \{(1, 4), (1, 5), (1, 6), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6)\}$
   d. $V = \{1, 2, ..., 6\}$, $E = \{(1, 4), (1, 5), (1, 6), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5)\}$

If you’re not convinced by the artistic lure of planar graphs, let’s show that they also have nice quantitative qualities, too. One cool thing that happens in a planar drawing is that the edges and vertices divide the plane into connected **faces**, much like the face of a cube. Note that we count the region outside of the boundary of the graph as a face too. Our favorite example graph has 4 faces.

4. In my graph and each of the graphs from problem 3 that you identified to be planar, count the number of vertices ($v$), edges ($e$), and faces ($f$). Use this data (along with data from other planar graphs you might want to draw) to come up with a formula relating these quantities. (For a hint, see the end of this Summer Fun problem set on the next page.)

5. Prove your formula for all connected, planar graphs! (For a hint, see the end of this Summer Fun problem set on the next page.)
Beyond the Plane

While the plane is our most readily-available graph-drawing surface, it certainly isn’t our only option. What’s stopping you from grabbing a sharpie and drawing one on the side of a baseball or mug? Of course, this can be hard to picture without the surface itself. But we can model many surfaces by taking a polygon and gluing its sides together.

The torus (better known as the donut) can be constructed by taking a square and gluing both pairs of opposite sides together (without twists). We can draw a graph on the torus by drawing a graph on the square and making sure the edges and vertices of the graph line up when we glue it together. Suddenly we find that graph drawings that weren’t possible in the plane are now possible on the torus!

6. Draw the graph from Problem 3b on the torus so that no edges cross each other.

The Möbius strip is a slightly stranger surface that we obtain by gluing two opposite edges of a rectangle after we have twisted the paper by a half turn (if you haven’t done this in person before, try it with a long, thin, strip of paper and a piece of tape). We denote this by putting the arrows on these opposite edges in opposite directions. Again, make sure edges and vertices on the boundary line up after the twist!

7. Draw the graph from Problem 3c on the Möbius strip so that no edges cross each other.

8. Count the vertices, edges, and faces in your drawings from Problems 6 and 7 (remembering that faces extend across the gluing boundary!). Do these surfaces obey the formula you found for Problem 4?

Hint to Problem 4: Can you fill in the blanks in the following formula with operations on the left side and an integer on the right side?

$$v \_ \_ \_ e \_ \_ \_ f = \_ \_ \_$$

Hint to Problem 5: What happens when you remove an edge?
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 12
Mentors: Jade Buckwalter, Anne Gvozdjak, Clarise Han, Shauna Kwag, Padmasini Krishnan, Hanna Mularczyk, Kate Pearce, AnaMaria Perez, Laura Pierson, Swathi Senthil, Jane Wang, Rebecca Whitman, Muskan Yadav, Valerie Yuen

May 4, 2023

We held our traditional end-of-session Math Collaboration. Math Collaborations are like escape rooms except that participants must solve lots of math problems to solve the event. For our end-of-session Math Collaborations, the math problems involve mathematics members have explored during the session. Starting with Alexandra Fehnel’s cubic Math Collaboration from Session 21, mentors have been dreaming up amazing plots and meta puzzles for this event (including one member-created event at the end of Session 30!). This session’s exciting Math Collaboration was created by mentors Elisabeth Bullock and AnaMaria Perez, who encoded the directions to a directional combination lock as a series of arrows hidden in mystery nonograms whose row and column specifications could only be obtained by solving math problems. Try your hand at solving a few of the problems from the event!

Using exactly 4 slices, cut a cake into
   a) 11 pieces.
   b) 14 pieces.
   c) 15 pieces.
   d) Using 2 slices, cut a donut into 5 pieces.
   e) Using 3 slices, cut a donut into 9 pieces.

What is the shortest path from the point \( p = (3, 2) \) that touches all the sides of the square with vertices (0, 0), (4, 0), (4, 4), and (0, 4), and then returns to \( p \)?

Karia has a grid of 7 hexagonal telescope cells and she wants to place them on a circular mounting board of minimal area. If the cells are identical with apothem of length 4, what is the minimal area of the mounting board?

What is the distance between the two common points of intersection of the graphs of \((x + 4)^2 + (y – 3)^2 = 126/4 \) and \(x^2 + y^2 = 25/4\)?

Each regular hexagon in the figure has side length 1. What is the area of the rectangle?
## Calendar

### Session 32: (all dates in 2023)

<table>
<thead>
<tr>
<th>Month</th>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>26</td>
<td>Start of the thirty-second session!</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>No meet</td>
</tr>
<tr>
<td>February</td>
<td></td>
<td></td>
</tr>
<tr>
<td>March</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>No meet</td>
</tr>
<tr>
<td>April</td>
<td>6</td>
<td>Karia Dibert, University of Chicago</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>No meet</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

### Session 33: (all dates in 2023)

<table>
<thead>
<tr>
<th>Month</th>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>September</td>
<td>14</td>
<td>Start of the thirty-third session!</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>October</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>November</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>Thanksgiving - No meet</td>
</tr>
<tr>
<td>December</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

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](http://www.girlsangle.org/page/math_collaborations.html).

Girls’ Angle can offer custom math classes over the internet for small groups on a wide range of topics. Please inquire for pricing and possibilities. Email: girlsangle@gmail.com.
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.

© Copyright 2023 Girls’ Angle. All Rights Reserved.
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:
- Connie Chow, founder and director of the Exploratory
- Yaim Cooper, Institute for Advanced Study
- Julia Elisenda Grigsby, professor of mathematics, Boston College
- Kay Kirkpatrick, associate professor of mathematics, University of Illinois at Urbana-Champaign
- Grace Lyo, assistant dean and director teaching & learning, Stanford University
- Lauren McGough, postdoctoral fellow, University of Chicago
- Mia Minnes, SEW assistant professor of mathematics, UC San Diego
- Beth O’Sullivan, co-founder of Science Club for Girls.
- Elissa Ozanne, associate professor, University of Utah School of Medicine
- Kathy Paur, Kiva Systems
- Bjorn Poonen, professor of mathematics, MIT
- Liz Simon, graduate student, MIT
- Gigliola Staffilani, professor of mathematics, MIT
- Bianca Viray, associate professor, University of Washington
- Karen Willcox, Director, Oden Institute for Computational Engineering and Sciences, UT Austin
- Lauren Williams, professor of mathematics, Harvard University

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: ______________________

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.

□ Enclosed is $216 for one session (12 meets)

□ I will pay on a per meet basis at $20/meet.

Nonmembers: Please choose one.

□ I will pay on a per meet basis at $30/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: ___________________________________________