WXML Final Report: OEIS/Wikipedia Task Force

Matthew Conroy, Travis Scholl, Stephanie Anderson, Shujing Lyu, Daria Micovic

Spring 2016

1 Introduction

Students worked to improve entries in the Online Encyclopedia of Integer Sequences (OEIS) and Wikipedia. Both are valuable resources for the mathematics community, and the world at large.

The OEIS began as the EIS in 1964, and has grown to over 270000 entries, each describing an integer sequence. Many entries list fewer than 50 terms. The encyclopedia allows the addition of so-called “b-files” which extend the number of the terms available. Adding b-files to OEIS entries was the primary way students contributed to the OEIS this quarter. Students gained experience contributing to the mathematics community while working at the standards level of a mathematical reference work. Meanwhile, they were exposed to a number of topics not necessarily part of a math undergraduate's normal education.

Wikipedia contains a vast collection of information on mathematics. Many mathematics articles, however, are quite minimal, lacking full descriptions, links to related topics, useful illustrations, etc. The task force focused this quarter primarily on adding informative, often animated, graphics to Wikipedia articles. Students gained skilled in graphics creation and design, while working with many mathematical concepts for the first time.

2 Contributions to the OEIS

A018224 This sequence is represented by $a(n) = \binom{n}{\lfloor \frac{n}{2} \rfloor}^2 = \text{A001405}(n)^2$. I approached this sequence in the easiest way possible. Since it takes each element of A001405 and squares it, I just took that sequence, which had more terms in the OEIS, and squared it to extend A018224. I increased the length of this sequence from 23 to 200. (SA)

A020001 This sequence represents the nearest integers to $\dfrac{\Gamma(n+\frac{11}{12})}{\Gamma(\frac{11}{12})}$. It is very similar to A020005 (Nearest integer to $\dfrac{\Gamma(n+\frac{10}{11})}{\Gamma(\frac{10}{11})}$) . I was introduced to the Gamma function for the first time. Rounding the results to the nearest integers, I extended both sequences to the first 500 terms to make sure the 1,000-digit limitation for each term is satisfied. (SL)

A061076 The definition of this sequence is $a(n)$ is the sum of the products of the digits of $1, 2, \ldots, n$. Note that for a number that ends in 0 the corresponding integer in the sequence is the same as the corresponding integer to the previous number. (eg. $a(9) = 45, a(10) = 45$ since $1 + 2 + \cdots + 9 = 1 + 2 + \cdots + 9 + 1 \times 0$). I extended this sequence to 10,000 terms. (DM)

A064572 The definition of this sequence is the number of ways to partition n into parts which are all powers of some integer k, where the exponents are not all 0. With the help of Travis, I used recursive method to increase the length of this sequence by uploading a b-file with 220 terms. The title used to be "Number of partitions into powers of integers greater than one". I revised the title and added comments to specify the restriction that the exponents cannot all be 0. (SL)

A064632 The definition of this sequence is the smallest prime p such that

$$n = (p - 1)/(q - 1)$$

for some prime q. This is a sequence related to the conjecture of Schinzel where if $f_1(x), \ldots, f_s(x)$ are irreducible polynomials with integer coefficients such that no integer $n > 1$ divides $f_1(x), \ldots, f_s(x)$ for all integers x, then there should exist infinitely many x such that $f_1(x), \ldots, f_s(x)$ are simultaneously prime [4]. Matthew M. Conroy verified to 10^9 that if this conjecture is true, every positive integer n can be represented as

$$n = (p + 1)/(q + 1)$$

with p and q prime [1]. Though these sequences are not the same, it is interesting to note that the OEIS includes sequences that related to other important sequences that could later be used to verify other conjectures in mathematics. Originally the definition did not include "for some prime q" and it had instead "and $p > q$." If we don't specify that q must also be prime we get a different sequence so we updated that definition. A b-file was added with 10,000 terms created using SageMath. (DM, SA)

A064652 This is the corresponding list of qs that go along with the A064632 sequence. So this is all primes q such that $n = \frac{p-1}{q-1}$ for the smallest prime p. Computing this sequence required me to return a different variable from the same code that I used to generate A064632. The b-file was extended from 100 to 1155 terms.(SA)

A075584 The definition of this sequence is the primes p such that the number of distinct prime divisors of all composite numbers between p and the next prime is 4. From the beginning it was unclear whether the definition of this sequence meant that each composite number between p and the next prime had 4 distinct prime divisors or that among all of the composite numbers, there were 4 distinct prime divisors. The second interpretation was correct and I wanted to update this distinction to the OEIS. However, my request to change the definition to "among all composite numbers" from "all composite numbers" was denied by the editors of the sequence. This was a lesson of mathematical wording preference. I extended this sequence to 10,000 terms. (DM)

A075587 The definition of this sequence is the primes p such that the number of distinct prime divisors of all composite numbers between p and the next prime is 7. This sequence is very similar to A075584. I incurred the same issue with the wording that was kindly corrected by the OEIS community. I extended this sequence to 10,000 terms. (DM)

A156681 This sequence stands for the ordered long legs of Pythagorean triangles, sorted to correspond to increasing short legs. I developed interest in this sequence when working on the scatterplot of the Pythagorean triples. I used Euclid's formula to generate the first 6,000 Pythagorean triples, and uploaded a b-file of the first 5,000 terms in correspondence with the existing short leg sequence A009004 which has 5,000 terms. (SL)

A257667 This sequence represents primes containing a digit 5, which is very similar to A257668 (Primes containing a digit 7). Using SageMath built-in functions such as primes and digits, I extended both sequences to the first 100,000 terms. (SL)

3 Contributions to Wikipedia

Bifolium At first, this Wikipedia page only included the quartic plane and polar equations. This was not enough information to understand the mechanics of the curve. I wanted to add a description for the construction of the curve like the one that was included on the Wikipedia page for the Cissoid. Knowing the construction of curves used to be necessary before the invention of graphing calculators. However, understanding the construction now is important because it helps with visualizing certain symmetries and other ways of thinking about how curves are made and how their equations are derived. I researched the bifolium and found an article [2] by Stephen Kokoska that described the construction of the curve. I modified and clarified the description of the construction and added it onto the Wikipedia page. My modification elaborated on how the lines extending were oriented. My updated definition was as follows: consider the circle C through the point O, and tangent line L to the circle at point O. For each point Q on C, draw the points P such that PQ is parallel to the tangent line L and $PQ = OQ$. The collection of points P forms the bifolium. (DM)

Cissoid of Diocles The Wikipedia page for this curve included the construction but not an animation to visualize that construction. I animated the construction to add intuition to the definition. Throughout the animation's various stages there was discussion among our group members about what was more effective in Wikipedia animations, aesthetics or precise graphical representations. I chose to adopt a balance of the two where the animation includes labels so the interaction of the components of the construction are understood, however I omitted the axes to give the animation a cleaner and more readable look. (DM)

Cochleoid A cochleoid is a snail-shaped curve represented by the polar equation r=a*sin(theta)/theta. The graphic on the page before was an svg but it was not created with enough points to make it very smooth, the aspect ratio was not 1 to present the true shape, and it only showed positive theta values. To improve the graphic, I created a polar plot in Sage using the equation $r = \frac{\sin\theta}{\theta}$ with $-20 < \theta < 20$. It is an svg with enough points, aspect ratio equal to 1 and both positive and negative ranges. The finished graphic has a beautiful symmetric heart shape. (SL)

Euclidean Algorithm Running Time This graphic represents the number of iterations used by the Euclidean algorithm to compute gcd(x,y), with red representing 0 and the following colors in the spectrum corresponding to higher values. I was encouraged to achieve higher resolution and experiment with the color scheme. I first implemented a counter to count the number of iterations used by the Euclidean algorithm to compute the gcd(x,y) for x and y from 0 to 500, and then used the handy matrix_plot command to generate the color map by setting cmap to a specific label. I spent a lot of time trying to find a color scheme that has a wide enough range and agrees with our intuition that the larger numbers are higher on the spectrum (darker) compared to smaller ones. Eventually, I was able to use cmap='rainbow_r' to satisfy all the requirements. Looking at the graph, we found that the Fibonacci numbers are the most computationally expensive using Euclidean algorithm as the darkest area in the graph follows the line whose slope is the golden ratio. In addition, the plot is supposed to be symmetric along the line y=x except that the upper half is generally one step more than the lower half. (SL)

Pythagorean Triples This graphic was created to show the parabolic patterns of the legs(a,b) of the Pythagorean triples. Originally, there were unnecessary labels and arbitrary boundary within a certain radius of the origin. I also planned to make an svg to be able to zoom in. As I started to work on perfecting the graphic, I first attempted to check if I could use two OEIS sequences as my data set. Unfortunately, the sequence A156681 (ordered long legs) had only fewer than 100 terms and I had to generate all the data from scratch. Assuming a¡b, I used the SageMath command list_plot to generate the scatterplot, with all the (a,b)s and (b,a)s reflected to both x- and y-axis. The arbitrary circular boundary and labels were removed. However, I still chose my final plot to be a png instead of an svg, considering that it was unnecessary to look at specific points, and that it was fairly computationally expensive to calculate each point every time we pull out the graph. (SL)

Watt's Curve This curve is created by taking two circles and creating a line segment whose ends are connected to each of the circles and as each end of the line segment moves about the circles the midpoint of the line segment creates Watt's Curve. In making this animation, I first had to learn about how to animate a graphic. The method I used was making many image files and then linking all the frames together. To make the curve, I used the polar equation (from J.W. Rutter's book *Geometry of Curves* [3]) and then between each frame had the midpoint of the line segment move about that curve. To make the whole line segment, I had to parametrize the line not only based on the polar equation, but also on the angle of the line segment such that each end stayed connected to the circle. Plus, because the polar equation is given as r^2, part way through the animation I had to switch to using the negative r, and then switch back, so that the animation stayed smooth. The shape can take on various forms depending on the radius of the circles, the distance between the circles, and the length of the line segment, so I have added to the Wikipedia page three animations, all of which take on various shapes of the curve. These animations will help give those interested a better understanding of all that Watt's Curve entails. (SA)

Wittgensteins Rod Wittgenstein's Rod is a geometry problem discussed by the 20th century philosopher Ludwig Wittgenstein. A ray is drawn with its origin 'A' on a circle, through an external point S and a point B is chosen at some constant distance from the starting end of the ray, and the shape of the figure that B describes when all the initial points on the circle are considered depends on three parameters: the radius of the circle, the distance from the center to S and the length of the segment AB. Originally, there was no equation or animation on its Wikipedia page. I created several animations with varying distances between the fixed point and the circle and varying length of the rod to present different shapes of the figure. The animations were generated by using the animate command on a bunch of frames, of which each frame represents a specific static plot of the system. To make each plot look smooth, I increased the number of frames as the distance between the fixed point and the circle decreases. The four animations I uploaded to Wikipedia present evolving patterns of the figure, from teardrop-shaped, to crescent-shaped, and eventually to bean-shaped as the fixed point moves inside the circle. (SL)

References

[1] M. M. Conroy. *A Sequence Related to a Conjecture of Schinzel*. J. Integer Sequences 4, No. 01.1.7, 2001. http://www.cs.uwaterloo.ca/journals/JIS/VOL4/CONROY/conroy.html

[2] Stephen Kokoska. *Fifty Famous Curves, Lots of Calculus Questions, And a Few Answers*. Department of Mathematics, Computer Science, and Statistics,Bloomsburg University, Bloomsburg, Pennsylvania.

[3] John W. Rutter. *Geometry of Curves*. Boca Raton, FL: Chapman & Hall/CRC, 2000.

[4] Eric W. Weisstein. *Schinzel's Hypothesis*. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/SchinzelsHypothesis.html