??, ??, look-say and rats

Supplies	Notes
To copy: Nothing Supplies: ● Whiteboard and markers, ● scratch paper (graph paper slightly prefered), pencils.	Author: Peter Tingley Email address: peter.tingley@gmail.com Dates taught: Resources: http://www.njohnston.ca/2010/10/a-derivation-of-conways-degree-71-look-and-say-polynomial/

Transferable Heuristics	● Organizing complex/confusing information ● Investigating novel structures. ● Understanding the value of multiple solutions ● Finding and celebrating unexpected connections.
Objectives:	● Students will generate interesting mathematical questions about the look-say sequence, and answer some of them. ● Students will develop organizational schemes to allow them to answer the rats question computationally. ● Students will understand the connection with the Fibonacci sequence (or with recursion in general, if they know about solving more general recursive sequences), and how to algebraically solve the problem. Ideally some of them will generate these connections on their own.
Teacher Overview	Students will begin by considering the look-say sequence. Put it on the board and let them go. Then get them to generate (and maybe answer) mathematical questions about it. The rat question is really two questions: first students need to organize information well enough to solve it computationally, which takes time. Then there is the question of finding closed form solutions or approximations, which also takes a lot of time. Allow students plenty of time for the initial organizational/brute force methods, and to understand the

	approximation techniques. It is fine to just show them the more exact answers at the end. Make sure you have 5-10min left to show the connection to the original look-see question. I think it is wonderful, and ties things together, but is too complicated to spend much time on.

Narrative

Time/Description	Activity (include links to any handouts)	Teacher Notes
1:00-1:10	Write about 10 lines of the look-say sequence on the board (see below), and ask students to (a) try to figure out what the sequence is, and (b) try to ask some interesting questions about it.	Ask students not to tell the answers to others when they figure it out. Give hints over the course of the 10 min, as needed. Good hints are "you will be angry at me", "it is barely math". Even more effective is to get someone to read it out loud to others, but save that until near the end of the time. If possible, put this in a place on the board where it can stay for the whole class, and leave room at the bottom for more rows.
1:10-1:30 (be flexible on end time)	Discussion of the look-say sequence. First make sure the whole class gets what it is. Tell them this sequence was introduced by John Conway (of the game of life), and he actually spent time studying it. So, it is silly, but smart people have been interested in it! Ask them to come up with interesting math questions to ask about it, and write them on the board. They will probably answer some of them.	Questions you should expect: 1) Does 4 ever show up? Higher digits? (no, see below) 2) Does 333 ever show up (no) 3) What if you start with a different digit? A digit greater than 3? A different string? (if start with a digit >3, looks the same whatever it is, and every line is a string of 1,2,3, with the original digit on the right) 4) Is the length of every line even after the first one? (yes) 5) Does the length always (weakly) increase? (Yes, non-trivial but I've seen people come up with correct arguments) 6) Are there always more 1s than 3s? (almost certainly yes) More generally, what ratios of digits do you expect? 7) Write down the length of each row. How does this sequence grow? (it is roughly exponential with base 1.3...but this is very non-trivial! See below) 8) Are there other (harder/simpler) versions of this question? I got this, but probably because I've trained people to always ask this question. Anyway, some very interesting answers are (a) do the same thing in

		binary (so sequence is 1,11,101,111011,11110101,100110111011...) (b) do the same thing with Roman numerals (I,II,III,IIII, IVI, IIIVII...) Both of these are a bit simpler because only 2 symbols appear...but are still too complex to solve in 2 hours! There is also the "unary" version: 1,11,111,1111,11111,... which is too easy, and doesn't give insight. Make sure some version of 7 gets on the board. Also, make sure some answer to some version of 1 ends up on the board (see below for some explanations). I once used this as a warmup for a different problem solving activity, and people spent the whole 2 hours just on it! It was an incredibly interesting discussion. So let this go on if it is going well!
1:30-1:55	Break students into groups of about 3. Write the following question on the board. Ask them to start working on it: A pair of young rats sneak onto a boat. They start reproducing. Assume that (on average) rats live 3 months, and females have -two children at the end of 1 month. -six more at the one of the second month -two more at the end of the third month. How many rats will there be after a year if no one deals with the problem?	One question that may come up is that the question doesn't state how many of the rats born each time are female. You'll have to get them to assume it is half, and may want to talk about how this means we will get an expected number of rats, not an exact number. For instance, there is some chance (1/1024) that all the offspring of the initial pair are male, in which case there are no rats in the end. There may also be questions about how old the initial rats are, when the first offspring appear... . Let them just make a reasonable choice. It will only affect the answer slightly. By the break, make sure everyone has realized they need to organize information carefully. It helps to realize that counting only female rats is a good idea, but don't force this.
1:55 - 2:05		

Break		
2:05-2:35	Groups work on their solutions to rat problem.	Keep encouraging good organization, but keep track of interesting schemes, even if inefficient. As time gets near the end, get students to put some of these on the board. If a group is super fast about the first part, try to get them to simplify until they get the recursion below. If anyone ever says "this is a bit like the Fibbonic sequence/Pascal's triangle," let them go with it! I actually don't know how, but last time I taught this people hit on very early. As students solve the question, ask them what would happen after 5 years (or more)...assuming a very big boat. I like to get them to approximate with an exponential, and lead people to this idea by pointing out that it is a simplified but basically realistic model for the growth of a population, and asking them what kind of growth they should expect (someone in each group should say exponential). I like to encourage people to pay attention to what they know about the real world! Once they approximate, ask them to try and figure out more precisely what the base is. A good hint here is "pretend it truly is exponential, and see what you can say about r".
2:35-2:50	Whole class discussion of rats question and various solutions	Get students to explain their organizational methods, ending with the 1-line recursion (if someone hit on that). These should already be on the board. Then discuss both the estimate and really good estimate expressions below. You can discuss the exact solution if you like, but I'm not planning to do much detail on that.
2:50-3:00	Briefly discuss why the growth of the look-say sequence is also controlled by an (incredibly complicated) recursion.	See the explanation below. This is show-and-tell, not precise math. Try to make the students see the beautiful hidden structure, but they will probably only understand on a vague level.

Proofs of some facts about look-say sequence

Fact: Every row after the first is even length

Proof: each row reads the previous row, so the digits come in pairs. (one two, or three one….)

Fact: No number greater than 4 appears

Proof: If N>4 appeared for the first time, there would be that many of the same number on the previous line. So, for example, there might be four 1s in a row:

...1111…

Then think about the line before that. Depending on parity, one of the following happens:

i) Reading the previous line, this says
… “one one, one one”...

But then the previous line would have been ….1 1….

And this would actually be read “two one”, which is a contradiction.

ii) There parity is oposite, so you need to consider the digit before these ones, so write it as ….x1111y… this reads

x one’s , one one, one y….

But then you should have actually read (x+1) ones, one y, so this is again a contradiction.

I explain this pretty informally by example...

Fact: 333 does not appear

Proof: An argument like the previous one shows that, if 333 is on one line, it must be on the line before as well, so this is true by induction!

Some “organizational” solutions. You may get others!

The basic observation you need to make is that, at each year, you need to not only know how many rats there are, but how old they all are. The nicest way I know is with a table. Hopefully I've got all the entries right...

Number of female rats of each age. Row enumerate months, with first row being the situation just before the first babies are born.

1 month	2 months	3 months	Total	0.4048*(2.413)^t
1	0	0	1	0.962
1	1	0	2	2.322
4	1	1	6	5.603
8	4	1	13	13.520
21	8	4	33	32.624
49	21	8	78	78.722
120	49	21	190	189.958
288	120	49	457	458.367
697	288	120	1105	1106.041
1681	697	288	2666	2668.876
4060	1681	697	6438	6439.997
9800	4060	1681	15541	15539.714

So, the number after 1 year is 2*15541=31082.

Another version of this that came up once is essentially the same, but with different notation. Instead of the table. You have
1_1
1_1,1_2
4_1,1_2,1_3
8_1,4_2,1_3
21_1,8_2,4_3
And so on. You keep track of the number of rats of each age, with the age recorded as a subscript. This is no different mathematically, but indicated

a different way of getting there.

Another super interesting organization I've seen is the following table, which keeps track of all rats along with the year they were born (which row they are on), and the generation they are part of (i.e. how many female ancesters lived on the boat) (column they are in):

1						
	1					
	3	1				
	1	6	1			
		11	9	1		
		6	30	12	1	
		1	45	58	...	
			30	144		
			9	195		
			1	144		
				58		
				12		
				1		

Then the number of females born in year k is the sum of the numbers on row k. To get number of females alive in year k, you have to sum all the numbers in that row and the two previous ones...so this is not trivial. But you can see some neat things: for instance, the sums of the columns are powers of 5 (why?). It also does look a bit Pascal's triangle-y: each entry is calculated as x+3y+z where x,y,z are the three entries in the column to the left starting on the row above it.

The cleanest way is to notice that columns 2 are 3 of the first table are just copies of column 1, but shifted down. So, you don't need to record them! To understand the number of 1 month old female rats, you have a sequence where, to generate the next term, you add the previous term, three times the term before that, and the term before that. This gives a recursive sequence (my students liked to call it a "modified Fibonacci sequence"):

1,4,8,21,49,120,288,697,1681,4060,9800,...

Defined by $a_n= a_{n-1}+3*a_{n-2} + a_{n-3}$.

Approximating by a geometric sequence

People should believe from their knowledge of biology that the growth will be exponential, so look like $c\ r^t$ for some c and r. The question is, what are c and r? To find r, just take the ratios of any two terms (far enough in)! For example, 15541/6438= 2.4139. Or you could use 6438/2666= 2.4128. See how similar they are.in practice, it doesn't matter which you use! So, the number of rats after t months is basically $c\ 2.413^t$ for some c.

To find c, also, just plug something in. For instance, t=12 should give $c*2.413^{12}= 15541$, and solving c=0.3988. But c is not so important… . You can see in the table how good this approximation is!

If you did this for just the number of rats born that year, instead of the total, c actually works out to exactly ¼! Reason is (sort of) explained below.

Approximating really well by a geometric sequence

The real source of error in the previous estimate is that we approximated r (since the errors grows as you take high powers). So, how to find r exactly? Pretend it really is geometric! So, the sequence starts

cr, cr^2, cr^3, cr^4

But, we also know that $a_4=a_3+3a_2+a_1$

So, $cr^4=cr^3+3cr^2+cr$

Solving (it is a cubic, but factors as a linear times a quadratic), $c=-1$, $1-\sqrt{2}$, or $1+\sqrt{2}$.

Since the sequence obviously grows, the answer must be $r=1+\sqrt{2}$, which is in fact approximately 2.41. So, the really good approximation is

$c\ (1+\sqrt{2})^t$,

And again you can solve and find c is about 0.398833.

Exact formula

The exact formula for the number of female rats born in month t is $(1+\sqrt{2})^t/4- (1-\sqrt{2})^t/4 - (-1)^t/2$

Don't derive this, but it may be worth writing down if there is time. The thing to notice is that the other two roots of the cubic make an appearence!

Explanation of look-say and connection to rats:

State: growth of row lengths is exponential, roughly 1.3^n. Actually r^n were r is a root of an irreducible degree 71 polynomial! How you ever seen degree 71 polynomial before??!!!!?!!!??

Show the hand-drawn version of the sequence, with the breaks marked, or draw it on the board (you can leave off the last 3 lines, but not more).

Say that there are places where the sequence breaks, and from then on acts as two disjoint triangles. This takes some explanation.

Say that, if you draw in all the breaks, there are only 92 different "words" that can show up between the cuts (after row 7).

So, to go from one row to the next, each "word" gives "birth" to various other words, and then "dies."

So, it is kind of like reproduction of animals, but there are 92 types of animals, and reproduction is...complicated.

Roughly, gives a recursion where you need to look 92 steps back! (this is not quite accurate; it is a Markov chain with a 92 by 92 matrix, which is even more complicated, but I'm ok with this tiny lie)

Solving like Rats gives a degree 92 polynomial...but it factors, and the growth rate is actually controlled only by a degree 71 polynomial!

Suggested Extension Questions

- Read about look-say sequence! It is a fun read, if nothing else.

Teacher Reflections

- When you reflect on how a session went for the benefit of future teachers, write it down here! Use comments for things that are ephemeral--that will get resolved and disappear.

1
11
21
1211
111221
312211
13112221
1113213211
31131211131221
13211311123113112211
11131221133112132113212221
3113112221232112111312211312113211
1321132132111213122112311311222113111221131221
11131221131211131231121113112221121321132132211331222113112211
311311222113111231131112132112311321322112111312211312111322212311322113212221
132113213221133112132113311211131221121321131211132221123113112221131112311332111213211322211312113211