Problem of the Month

Ian VanderBurgh

Next in the sequence of Problems of the Month comes a problem (actually, two problems) about sequences.

Problem 1 (2006 Pascal Contest)

John writes a number with 2187 digits on the blackboard, each digit being a 1 or a 2. Judith creates a new number from John's number by reading his number from left to right and wherever she sees a 1, writing 112, and wherever she sees a 2, writing 111. (For example, if John's number begins 2112, then Judith's number would begin 111112112111.) After Judith finishes writing her number, she notices that the left-most 2187 digits in her number and in John's number are the same. How many times do five 1s occur consecutively in John's number?

This is quite something to try to wrap your head around! Let's start off with a slightly different problem, which is a bit easier to sort out:

Problem 2. John writes a number with 3000 digits on the blackboard, each digit being a 1 or a 2. Judith creates a new number from John's number by reading his number from left to right and wherever she sees a 1, writing 122, and wherever she sees a 2, writing 111. (For example, if John's number begins 2112, then Judith's number would begin 111122122111.) After Judith finishes writing her number, she notices that the left-most 2500 digits in her number and in John's number are the same. What are the digits in positions 2006 to 2012? (Problem 2 appeared on one of the earlier drafts of the 2006 Pascal Contest, but was removed in favour of Problem 1.)

It is tough to know where to start this problem; as the song says, though, let's start at the very beginning (it's a very good place to start).

Solution to Problem 2: If John's number begins with a 1, then Judith's will begin 122; if John's number begins with a 2, Judith's will begin 111. Since Judith's number begins with a 1 regardless of what John's number begins with, and since their numbers are the same for the first 2500 digits, John's number must begin with a 1. Thus, Judith's begins 122, and therefore John's begins 122. (We call these three digits Stage 1.)

We can then use this beginning as the "seed" with which to create the number. Since John's number begins 122, then Judith's begins 122111111 (since each 1 becomes 122 and each 2 becomes 111). Thus, John's number begins 122111111 (Stage 2).

Continuing in a similar way, each digit at each Stage generates three digits at the next Stage, which means that each Stage is three times as long as the previous one. Thus, successive stages consist of 3, 9, 27, 81, 243, 729, and 2187 digits. In general, Stage n consists of 3^n digits.

Consider Stage 7, which has length $3^7 = 2187$. This includes the digits in positions 2006 to 2012, the positions that interest us. How can we figure out what these digits are? One strategy is to trace backwards through the Stages the positions of the digits which generate these digits, until we come to one of the early Stages, where we can easily determine the digits. We then use these known digits to move forward again through the Stages to determine the desired digits at Stage 7.

In general, the first k digits from one Stage generate the first $3k$ digits at the next stage. This tells us that digit k from one Stage generates digits $3k-2$, $3k-1$, and $3k$ at the next stage. We would like digits 2006 to 2012 at Stage 7. Digit 669 from Stage 6 generates digits 2005 to 2007 at Stage 7, digit 670 generates digits 2008 to 2010 at Stage 7, and digit 671 generates digits 2011 to 2013 at Stage 7. If we can determine digits 669 to 671 at Stage 6, then we can get digits 2006 to 2012 at Stage 7. (These digits at Stage 6 actually will give us digits 2005 to 2013 at Stage 7, which is a bit more than we need.)

Perhaps a chart might be in order:

Stage	Digit(s)		Stage	Digit(s)		Digit(s)
7	2006–2012	from	6	669–671	giving	2005–2013
6	669–671	from	5	223–224	giving	667–672
5	223–224	from	4	75	giving	223–225
4	75	from	3	25	giving	73–75
3	25	from	2	9	giving	25–27

Aha! We know that digit number 9 at Stage 2 is a 1. Hence, remembering that a 1 at one stage gives 122 at the next stage and a 2 at one stage gives 111 at the next stage, we can then trace these digits forward through the Stages. Probably making another table will help:

Stage	Digit(s)	String		Stage	Digits	String
2	9	1	gives	3	25–27	122
3	25	1	gives	4	73–75	122
4	75	2	gives	5	223–225	111
5	223–224	11	gives	6	667–672	122122
6	669–671	212	gives	7	2005–2013	111122111

Thus, the digits in positions 2006 to 2012 are 1112211.

Making a table seemed to be pretty helpful here. Often, it is a really good way to consolidate or organize a bunch of information. Now Problem 1 certainly looks similar (be careful, though—the digit replacement scheme is slightly different).

Solution to Problem 1: Conveniently, here we are looking at 2187 digits, which exactly matches the length of Stage 7 in our analysis above. (This seems unlikely to be a coincidence!) Here, the digit replacement scheme is that 1 becomes 112 and 2 becomes 111. Again, John's number must start with a 1, and thus must start 112 (Stage 1).

How can 11111 (that is, five consecutive 1s) appear at Stage 7? They must come from 21 at Stage 6. They cannot come from 22 because there cannot be two consecutive 2s at any stage. (Think about why this is true.) In fact, there cannot be more than five consecutive 1s at any stage. (Think about this as well.)

Thus, counting the number of occurrences of 11111 at Stage 7 is the same as counting the number of occurrences of 21 at Stage 6. But each 2 is always followed by a 1; whence, this is the same as the number of 2s at Stage 6. (Wait! If the 2 occurs at the very end of Stage 6, then it is not followed by a 1. It turns out that Stage 6 ends in a 1—a third thing to think about!)

How can we figure out the number of 2s at Stage 6? The only way to get a 2 at Stage 6 is from a 1 at Stage 5, which gives 112. (And each 1 at Stage 5 gives exactly one 2 at Stage 6).

At this point, trying to write all of this out in words is becoming painful. Thus, some notation would be helpful. Let's use a_n to represent the number of 1s at Stage n and b_n to represent the number of 2s at Stage n. Notice that $a_1 = 2$ and $b_1 = 1$. Also, $b_{n+1} = a_n$ (each 1 at Stage n generates exactly one 2 at Stage $n+1$) and $a_{n+1} = 2a_n + 3b_n$ (each 1 at Stage n generates two 1s and each 2 generates three 1s). Again, a table:

n	a_n	b_n
1	2	1
2	7	2
3	20	7
4	61	20
5	182	61

Therefore, the number of 1s at Stage 5 is 182, which implies that there are 182 occurrences of 11111 at Stage 7 (that is, in the first 2187 digits of the full number).

Yes, this problem was pretty difficult to be on a Grade 9 Contest(!). But it required very little specialized knowledge, and lots could be discovered through some fiddling, which was why we felt comfortable including it. And there is a lot more that could be asked, and lots of interesting investigation to be done (including lots that could be investigated with a computer). Happy hunting! It is also interesting to look at what happens by modifying the digit replacement schemes, even allowing the 1 and 2 to be replaced by strings of different lengths.