Barcodes!

Felipe Voloch

These notes and the barcode program are available at

`http://www.ma.utexas.edu/users/voloch/barcodes.html`

Casting out nines

$$
\begin{array}{r}
36282 \\
74363 \\
62111 \\
60011 \\
+34948 \\
\hline
267715
\end{array}
$$

Casting out nines

36282	$\rightarrow 3+6+2+8+2=21$	$\rightarrow 3$
74363	$\rightarrow 7+4+3+6+3=23$	$\rightarrow 5$
62111	$\rightarrow 6+2+1+1+1=11$	$\rightarrow 2$
60011	$\rightarrow 6+0+0+1+1=8$	$\rightarrow 8$
+34948	$\rightarrow 3+4+9+4+8=28$	$\rightarrow 10 \rightarrow 1$
267715	$\rightarrow 2+6+7+7+1+5=28$	$\rightarrow 10 \rightarrow 1$
and	$3+5+2+8+1=19$	$\rightarrow 10 \rightarrow 1$

$$
\begin{array}{r}
3467 \\
\times 833 \\
\hline
2888011
\end{array}
$$

3467	$\to 3+4+6+7=20$	$\to 2$
$\times 833$	$\to 8+3+3=14$	$\to 5$
2888011	$\to 2+8+8+8+0+1+1=28$	$\to 10 \to 1$
and	$2 \times 5 = 10$	$10 \to 1$

The remainder of division of $a+b$ by 9 equals the sum of the remainders of division of a and b by 9.

The remainder of division of $a + b$ by 9 equals the **remainder** (upon division by 9) of the sum of the remainders of division of a and b by 9.

There is nothing magical about 9 there and you can replace it by any number you like.

UPC barcodes

First of all, every product maunfactured in the US, has a UPC code which consists of a 12-digit number, such as 022400004419. The first digit designates the region of manufacture, the next five the company that makes the product, next five the product code given by the company and the last one is a *check digit* which is chosen so that the sum of the digits in the even positions (i.e. second, fourth, ...) plus three times the sum of the digits in the odd positions (i.e. first, third, ...) is divisible by 10. For example, in the above example we compute

$$2 + 4 + 0 + 0 + 4 + 9 + 3 \times (0 + 2 + 0 + 0 + 4 + 1) = 40$$

The purpose of this is to detect errors. If just one digit is read incorrectly then the sum will not come out divisible by 10. To actually correct the error we need to know which digit was incorrectly read.

There is an extension of the UPC codes called EAN-13, that uses 13 digits and is used worldwide.

digit	left	right
0	0100111	1110010
1	0110011	1100110
2	0011011	1101100
3	0100001	1000010
4	0011101	1011100
5	0111001	1001110
6	0000101	1010000
7	0010001	1000100
8	0001001	1001000
9	0010111	1110100

The bars are made out of the strings of 1's. Every number is represented by two bars and two spaces of width 1,2 or 3 so that the sum of the lengths of the bars is even. The left numbers start with a space and end with a bar and the right numbers start with a bar and end with a space. You may amuse yourself showing that under these rules we cannot represent more than 10 digits using seven bits.

What good does it do to have left and right different?

What good does it do to have always an even number of bits?

Error correction

Here is a scheme for error-correction. The data is represented by a string of 11 digits such as 12746763710 such that both the sum of the digits **and** the expression $a_1 + 2 \times a_2 + 3 \times a_3 + \cdots$ (where $a_1 a_2 a_3 \ldots$ are the successive digits) are both divisible by 11. So in the above example we need to compute $1 + 2 + 7 + 4 + 6 + 7 + 6 + 3 + 7 + 1 + 0 = 44$ and $1 \times 1 + 2 \times 2 + 3 \times 7 + 4 \times 4 + 5 \times 6 + 6 \times 7 + 7 \times 6 + 8 \times 3 + 9 \times 7 + 10 \times 1 + 11 \times 0 = 253 = 23 \times 11$.

If, our data is read and just one error occurs we correct it by first computing the sum of the digits modulo 11, which will give us the magnitude of the error. Now the second sum will tell us where the error is, as follows. If the magnitude of the error was j, we look at the j-th row of the array below for the only place where the second sum appears. The column we are then tells us where the error occurred.

1	2	3	4	5	6	7	8	9	10
2	4	6	8	10	1	3	5	7	9
3	6	9	1	4	7	10	2	5	8
4	8	1	5	9	2	6	10	3	7
5	10	4	9	3	8	2	7	1	6
6	1	7	2	8	3	9	4	10	5
7	3	10	6	2	9	5	1	8	4
8	5	2	10	7	4	1	9	6	3
9	7	5	3	1	10	8	6	4	2
10	9	8	7	6	5	4	3	2	1

Here is an example: 76364324610. The sum of the digits leaves a remainder of 9 when divided by 11. The expression $a_1 + 2 \times a_2 + 3 \times a_3 + \cdots$ leaves a remainder of 2 when divided by 11. In the 9th row, the number 2 occurs in the 10th column. So the error is of 9 in the 10th place. The value on the 10th place is currently 1. What number plus 9 gives a remainder of 1 when divided by 11? The answer is 3, so replace 1 by 3 in the 10th place to get the correct data 76364324630.

Why does this work? The main reason is that the table (which has the remaiders by division by 11 of the 10×10 multiplication table) has all digits in every row and column. This happens because 11 is a prime number and that's why we are not using 9 or 10. We could even use an extra symbol (X) to stand for a digit of 10. Something like that is used in the ISBN code for books. For a complete explanation we would need to show that given the error magnitude e and location j we can recover e and j from e and the remainder of the division of $e \times j$ by 11. I'll let you do as much detail as you think is appropriate by yourselves.