Chinese math helps DNA research

A 2000 year old Chinese mathematical theorem is making a difference in DNA sequencing, says S.Ananthanarayanan.

The Chinese theorem first found application in solving the kind of problem that is posed by the popular mind bender called SUDOKU. The way it helped was that a whole lot of possible solutions could be clubbed into one into one lot with a unique identification. This enabled classification of possible solutions and economy in the coding that had to be done for solving the problem – only the lot had to be identified, not the individual solutions.

Scientists at Cold Spring Harbor Laboratory, New York, have used the same principles to greatly simplify the formidable task of solving puzzles of the structure of the million-unit long DNA molecule.

Chinese Remainder Theorem

This was first described by the Chinese mathematician Sun Tsu in the 4th century AD as a statement of the relation between numbers, divisors and remainders. A simple relation exists between numbers that give the same remainder when divided by some number. For example, 14 divided by 3 gives remainder 2, just like 17 divided by 3 also gives a remainder 2, or 38 divided by 3. We can see that the differences between the numbers, (17-14 =3), (38-14 = 24) or (38-17=21) are all multiples of the divisor, 3. Number of this kind are called '***congruent to the modulus of 3***'.

Yet another relation is when a number gives certain remainders when divided by different divisors, like the number 25, gives a remainder of 5 when divided by 5, a remainder of 1 when divided by 8 and a remainder of 7 when divided by 9.

Now, the Chinese remainder theorem says that we can always find a number, like 25, which gives specific remainders when divided by a set of divisors that have the property that 'no pair of them have a divisor other than 1'.

For example, of the numbers, 10,7,33,13, no pair has a common divisor other than 1. But this is not true of 10,7,33,14, because the pair (10,14) has a divisor, 2. The divisors, 5,8 and 9, in our example, satisfy this condition. The condition is called being '***pairwise coprime***'.

The Chinese theorem says that if we take any number, divide it successively by divisors like this and list out the remainders, then it is possible to work out the original number! When we list out remainders, what we are actually doing is to specify the list of 'congruent' numbers, which would produce the same remainder, for each divisor. The theorem then says that if the divisors are pairwise coprime, there would be a common

number in all the congruent lists, which would yield the right remainder with all the divisors.

Sudoku

In the game of Sudoku, we need to create series of nini numbers each of which satisfy conditions in a row, a column and also in a 3x3 square. Close parallels have been developed between these conditions and the conditions in the Chinese remainder theorem and the theorem has helped develop methods to solve the Sudoku puzzle or even to say whether the puzzle has more than one solution or no solution!. The appearance of prime numbers in the theorem (a collection of prime numbers must be pairwise coprime) even helps crack difficult codes based on the product of two large prime numbers, a code that is used in e-commerce. The game of Sudoku, like the Rubik cube, has set in motion much research in mathematics, which have application in different fields, like study of networks and scheduling.

DNA sequencing

The genetic identity of any living thing is a specified by a giant molecule in the nucleus of each cell living thing. It has been established that the coding is with the help of groups of just four kinds of chemical groups that are attached, in a series, to the backbone of the DNA molecule. Each of these groups, which are threesomes of the chemical groups, defines a specific building block, called amino acids, of proteins, and myriad proteins can be defined with different combinations of these twenty amino acids, in a string that is often millions of units long!

Now, although the structure and nature of DNA is known, getting into the details of actual sequences of triads that define amino acids is completely different story. The structure of the DNA is at the scale of atoms and there are no forceps that can hold different portions apart. The detailed structure has to be deduced by identifying small and invisible portions of the sequence and looking for clues of adjacent segments with the help of chemical markers and statistical methods – a virtual hunt for a needle in the haystack of millions of sections of the DNA in feverish Brownian motion.

The traditional method has been to identify individual portions of DNA, mark them in some way and then try to locate specific structures. The method has worked and much progress has been made, but it is painstaking and tedious and holds out no hope of rapid sequencing for practical use.

The New York scientist have used ideas of the Chinese remainder theorem to simplify the task of classifying and analyzing millions of items of data that DNA research throws up. The researchers can now deal with whole libraries of genetic information instead of looking at just "one genetic sequence at a time," says Yaniv Erlich, the lead author of the paper that has appeared as the cover story of this month's ***Genome Research.***

The method has tremendous potential for clinical applications and can be used, says Prof. Hannon, leader of the group that developed the technique, to analyze specific regions of the genomes of a large population and identify individuals who carry mutations that cause genetic diseases.

[the writer can be contacted at simplescience@gmail.com]

The remainder conundrum

The Chninese, Sun Tsu Suan-Ching (4th century AD) posed the problem like this:

There are certain things whose number is unknown. Repeatedly divided by 3, the remainder is 2; by 5 the remainder is 3; and by 7 the remainder is 2. What will be the number?

The Indian mathematician, Brahmagupta (born 598 AD) puts it like this in *Brahma-Sphuta-Siddhanta*:

A woman goes to market and a horse steps on her basket and crushes the eggs. The rider offers to pay for the damage and asks how many eggs she had. She does not remember the exact number, but when she had taken them out two at a time, there was one egg left. The same happened when she picked them out three, four, five, and six at a time, but when she took them seven at a time they came out even. What is the smallest number of eggs she could have had?

The Chinese remainder theorem helps answer these questions, which is to trace the number which is *simultaneously congruent* to the different numbers which give certain remainders when divided successively by *pairwise coprime* divisors

---

The simple remainder theorem

This high school horror can also be put simply. Suppose we have a collection of powers of some number, which we call, 'x', like this for example:

$x^2 - 3x + 2 = 0.$

The task is to find all the values of x which would satisfy the relation. The remainder theorem says that if x = p is a solution, then, dividing the expression by (x-p) will leave no remainder. This is another way of saying that the original expression is a multiple of (x-p). The answer after dividing by (x-p) is then also a divisor without a remainder and hence contains another solution to the original relation, and so on.

In the expression: $x^2 - 3x + 2 = 0$, we can see that x=1 is a solution, because putting x=1 makes the expression : 1 -3 + 2, which = 0.

Let us divide by (x-1). To get the first term, x^2 from (x-1), we need to multiply (x-1) by x, to yield : x^2- x. This gives us '-x' in addition to x^2.

So we add 'x', to compensate the '-x', and the relation becomes:
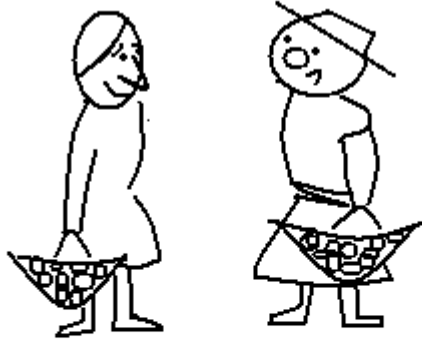
(x-1).x + x -3x +2 =0. This is the same as:

(x-1) -2x +2 =0, which can be written:

(x-1) – 2.(x-1) =0, which is the same as:

(x-1) . (x-2) = 0.

This gives us the second solution : x=2.

Eggs in basket problem:

2 children went to town to sell eggs. One had more in her basket than the other. She said, "Give me one egg and I will have twice as many as you." The other said, "Give me one of yours and we will have the same number." How many eggs did they have