Original Papers

Polish Psychological Bulletin
2009, vol. 40 (2), 69-73
DOI - 10.2478/s10059-009-0023-1

*Makoto Yamaguchi**

Savant syndrome and prime numbers

Oliver Sacks (1985) reported that a pair of autistic twins had extraordinary number abilities and that they spontaneously generated huge prime numbers. Such abilities could contradict our understanding of human abilities. Sacks' report attracted widespread attention, and several researchers speculated theoretically. Unfortunately, most of the explanations in the literature are wrong. Here a correct explanation on prime number identification is provided. Fermat's little theorem is implemented in spreadsheet. Also, twenty years after the report, questionable aspects were found in it. Extreme abilities became dubious. One possibility for the less extreme abilities is incomplete trial division.

Keywords: *autism, savant syndrome, exceptional abilities*

Introduction

Oliver Sacks (1985) reported that a pair of autistic twins, who were already famous for calendar calculation, also had extraordinary number abilities. He reported that they spontaneously generated huge prime numbers despite the fact that they lacked the abilities for even simple arithmetic. They were reported to have generated 6-digit primes at first. As Sacks challenged them with larger primes, they were reported to have identified 10-digit numbers as primes and generated even larger numbers, but Sacks could not confirm their primality as the range of number exceeded his prime number list. These reported abilities could contradict our understanding of human abilities and inevitably attracted widespread attention. However, in 2005, an important fact was shown by Yamaguchi (published as Yamaguchi, 2007a) and as a result researchers are forced to reconsider Sacks' report. This is explained later.

In response to Sacks' report, several researchers published theoretical speculations. Unfortunately, most are wrong. This article corrects errors in the literature (Sacks, 1985; Welling, 1994; White, 1988) and provides an exposition on prime number identification. Note that as researchers on prime numbers always require up-to-date information (e.g., largest know prime), the latest information is collected in an authoritative website, Prime Pages (http://primes.utm.edu/) run by Chris Caldwell of The University of Tennessee at Martin. Prime Pages also contains tutorials on prime number identification. As all the algorithms mentioned in this article are explained in Prime Pages, the reader can consult it for more rigorous details.

Does this report matter?

Do the twins' reported abilities pose challenges to scientific knowledge? Absolutely. Sacks rightly expressed his dismay at what he saw, because straightforward methods for checking whether certain numbers are prime require an enormous amount of computation to reproduce such abilities (see below for more details). For this reason, Dehaene (2001) expressed skepticism about such abilities. In addition, Sacks claimed that they could not calculate even small numbers correctly, and that they seemed not to understand division. Also it is important to note that the ability of another famous savant (Anderson, O'Connor, & Hermelin, 1998; Hermelin & O'Connor, 1990) is not as spectacular as one may consider at first. His abilities reported in Hermelin and O'connor (1990) are consistent with trial division by only 2, 3 and 5. Moreover, whether it can be divisible by 2 and 5 is easily seen by only looking at the last digit, leaving trial division by only 3. (Incidentally, reporting by Hermelin & O'Connor lacks rigor; in the data, the number of significant digits is inconsistent, and many are unnaturally clear-cut digits, e.g., *.00). The largest number used in Anderson et al.'s experiment is 993, whose

* **The University of Tokyo, Hongo, Japan; email: yamag-psy@toki.waseda.jp**

square root is about 32. To test primality, one needs to try to divide by only 11 primes, from which 2 and 5 may be excluded, leaving 9 primes. This range of ability is not a mystery. (Dehaene, 2001, already expressed the same view about him). The report of the twins by Sacks is the only one that poses challenges.

Before considering various methods for primality tests, one thing should be considered. Are all primes equal? Or are some easy to identify? Consider the same questions about composite numbers. It is easy to notice that composite numbers vary in their ease in being identified as composite. Even a very huge number, say of 100 digits, can be instantaneously concluded to be composite, if the last digit is even (or 5). In contrast, if that number is a product of two huge prime numbers, even a computer takes a long time finding the factors.

But how about prime numbers? Did the twins simply select easily identifiable primes? There do exist a special kind of prime numbers, which are called Mersenne primes. For numbers expressed in the form of 2^n-1, there exists an especially fast algorithm for primality tests. For this reason, most known largest primes are Mersenne primes. However, Mersenne primes are very rare. For instance, there are no 5-digit Mersenne primes, and there are only two 6-digit Mersenne primes. This implies that the primes generated by the twins were not especially easily identifiable primes, but usual primes, and that there were no shortcuts.

It is also worth noting that, according to mathematicians, it is unknown who first discovered the notion of prime numbers. If some savants are confirmed to have derived the notion without any book or education, this implies that "folk mathematics" can include the notion of prime numbers. This would be an important scientific fact.

Straightforward methods

Naively, to identify the primality of a certain number, one should try to divide that number by smaller primes (trial division). It was already pointed out that trial division and the sieve of Eratosthenes were confused in the literature (Yamaguchi, 2005). (For more details on both methods, see Yamaguchi, 2005, or Prime Pages). White (1988) claimed to propose a new algorithm, called "addition series", which is also reviewed in Welling (1994). However, it is nothing other than the sieve of Eratosthenes (more precisely, its suboptimal version). Also note that whereas trial division presupposes that one has a list of small prime numbers (otherwise, the required number of division increases), the sieve of Eratosthenes does not require any prior knowledge. However, the latter method requires a huge memory of a large number of integers. To test whether a certain number is prime, trial division is faster, but the sieve of Eratosthenes is faster when generating a list of prime numbers. Considering the nature of Sacks' challenge to the twins, trial division would have been the more appropriate method.

Sacks seems to be aware only of such straightforward methods. And he correctly suspected that such a feat with these methods would contradict our understanding of human abilities. Indeed, the requirement of memorization for already generated primes up to even 1,000,000 far exceeds the world record of reciting pi (Mr. Akira Haraguchi set the record at 83431 digits in 2005, then expanded it to just 100000 digits in 2006), which also precludes using the sieve of Eratosthenes. (However, as the calculation of long digits of pi was enabled relatively recently by computers, this is unlikely to be very close to the limit on human abilities). Trial division requires more than 150 divisions to confirm primality of the numbers around 1,000,000. It requires, for 100003 (the smallest 6-digit prime), 65 divisions. The 1000th prime is 7919, which implies that to test primality of numbers around 62,710,561 (7919 squared) requires 1000 trial divisions. Sacks' report goes far beyond, so the abilities are unlikely to be explained by trial division.

However, such straightforward methods are never used for large numbers. Before examining more practical methods, several caveats in reading Sacks should be seen. In his theoretical speculation, he cited Stewart (1975). This book is recommended reading for the basics of number theory, although it does not contain information on practical primality tests. Sacks misread Stewart and suggested the pigeon-hole principle could be used in primality tests for huge numbers. Actually, Stewart first saw a proposition holds for the first few prime numbers, but he said it might not generalize for other (larger) primes and sought to adopt a formal approach, using the pigeon-hole principle. It is required for formal proof, but has nothing to do with the size of the numbers. So the pigeon-hole principle is not relevant here. Also Sacks cited a letter from his colleague, in which division in modular arithmetic is mentioned. The reader must distinguish between division by 7 in usual arithmetic, which is relevant to calendar calculation, and division in modular arithmetic with mod 7. The calendar is naturally expressed in modular arithmetic with mod 7, in which division by 7 corresponds to division by 0 in usual arithmetic and is not allowed. Division in modular arithmetic may not be necessarily relevant here.

In the rest of this article, modular arithmetic is introduced. Readers unfamiliar with it are referred to an introductory textbook. (Stewart is recommended). One basic fact is that

if $a \equiv a' \pmod{n}$ and $b \equiv b' \pmod{n}$, then $a \cdot b \equiv a' \cdot b' \pmod{n}$

This directly implies that

if $a \equiv a' \pmod{n}$, then $a^x \equiv a'^x \pmod{n}$

This is used later in Fermat's theorem.

Practical methods

Practically used methods for primality tests are classified into two large categories; deterministic and probabilistic tests. Already mentioned methods are deterministic, so they never make any mistakes. An important starting point in understanding practical methods will be Fermat's (little) theorem, which is probabilistic. Stewart mentioned this as well as another theorem, Wilson's theorem (deterministic). Let us consider the two: for prime number p,

Fermat's theorem

$a^{p-1}\equiv 1 \pmod{p}$

Wilson's theorem

$(p-1)!\equiv -1 \pmod{p}$

As for Wilson's theorem, if n is prime then always $(n-1)!\equiv -1 \pmod{n}$, and conversely if $(n-1)!\equiv -1 \pmod{n}$, n is concluded to be prime. In contrast, attention is needed for Fermat's theorem. If n is prime then always $a^{n-1}\equiv 1 \pmod{n}$, but the converse does not hold. That is, $a^{n-1}\equiv 1 \pmod{n}$ may hold for some composite numbers. (i.e., the Fermat test fails for some composite numbers). In any event, if $a^{n-1}\not\equiv 1 \pmod{n}$ then n can always be correctly concluded to be composite.

The next thing to consider is whether these can be a practical test. At first sight, both seem to require computing extremely huge numbers. However, the situation is completely different. As computing factorials is often intractable even with the use of a computer, Wilson's theorem is only for theoretical significance. In contrast, although Fermat's theorem also seems to require intractable computation in terms of exponential, this can be reasonably handled.

To generate a huge number by exponential, consider the form $((((a^2)^2)^2)^2)^2$. This makes calculation of exponentials efficient. As a concrete example, let us compute $2^{200} \pmod{11}$. Literally, of course, this requires multiplying 199 times. The results, which have 61 digits, should be simplified for mod 11. However, the following method dramatically reduces the work:

$2^2\equiv 4$

$2^4=(2^2)^2\equiv 4^2=16\equiv 5$

$2^8=(2^4)^2\equiv 5^2=25\equiv 3$

$2^{16}=(2^8)^2\equiv 3^2\equiv 9$

$2^{32}=(2^{16})^2\equiv 9^2=81\equiv 4$

$2^{64}=(2^{32})^2\equiv 4^2=16\equiv 5$

$2^{128}=(2^{64})^2\equiv 5^2=25\equiv 3 \pmod{11}$

Using these results,

$2^{200}=(2^{128})\cdot(2^{64})\cdot(2^8)=3\cdot 5\cdot 3=15\cdot 3\equiv 4\cdot 3=12\equiv 1 \pmod{11}$

Notice that there was no need to handle huge numbers. (In this specific case, in the last line, simplifying $15\cdot 3\equiv 4\cdot 3$ would seem unnecessary. However, this operation proves useful for larger products. Even when many numbers are multiplied, an intermediate product can always be reduced to a number smaller than modulo). Therefore, exponentials can be made tractable in modular arithmetic, and Fermat's theorem can be used as a practical method for primality tests.

In the Appendix, Fermat's theorem is implemented as a spreadsheet, which should facilitate our understanding. For the Fermat test, first, base a should be selected. One can adopt a small number, and usually 2 suffices. Recall that if $a^{n-1}\not\equiv 1 \pmod{n}$, then the number is safely concluded to be composite. Even when it passes the Fermat test (i.e., $a^{n-1}\equiv 1 \pmod{n}$), the number may not be prime (hence called probabilistic method), so one should continue testing with different bases. Composite numbers that pass the Fermat test are called pseudoprimes. However, one should bear in mind an important fact: pseudoprimes are very rare. If one tests a number and it passes the Fermat test with base 2, then it is prime more than 99% of the time. In addition, the Fermat test can be made stronger repeated with different bases. That is, a pseudoprime with base 2 may be revealed as composite with repeated testing with a different base. This implies that, if the autistic twins somehow used the Fermat test with only one or a few bases, unless Oliver Sacks tested them very extensively using hundreds of numbers, they would not have erred by generating pseudoprimes. Although very exceptional composite numbers which always pass the repeated Fermat test exist, they are extremely rare. The Fermat test will be the strongest practically usable method available at that time. Ramachandran and Blakeslee (1998) proposed testing savants and seeing whether the same errors are made as probabilistic algorithms. It is very likely that Ramachandran meant the Fermat test.

More recent methods

Given the long history of number theory, one is tempted to think that there have been only minor developments in recent years. Therefore, it should be striking that significant progress has been made after Sacks' encounter with the twins. However, speculating that the twins somehow unconsciously derived such methods for the first time in the world would be too uncomfortably close to science fiction. For this reason, methods that were more recently developed are only briefly mentioned here.

First consider probabilistic methods. The Fermat test can be improved with only a slight additional effort. The Miller-Rabin test is so reliable that it is virtually indistinguishable from deterministic tests. It was developed after Sacks' encounter with the twins (but before the publication of his report). Even though its invention is relatively recent, it is not much more difficult or complex than the Fermat test.

In fact, there already existed strong deterministic algorithms when the report appeared. An example is the APR algorithm, which was developed after Sacks' encounter. Most recently, the historically important AKS algorithm was developed in 2002.

Is this report true?

Research on the abilities reported by Oliver Sacks saw a surprising turn in 2005. According to Sacks' report, he joined the twins with a book of a prime number list up to 10 digits. However, the number of primes up to 10 digits far exceeds what can be contained in a book. It is unclear whether his purported book contained all or some of the 10-digit primes. Even the most modest estimate is about 50,000,000, which cannot be included in a book. When questioned about it, Sacks answered that he no longer has the book or other resources of that time and that the book might have been up to 8 digits (see Yamaguchi, 2007a. Also see Snyder, 2007, for commentary, and Yamaguchi, 2007b, for author response. Incidentally, the author does not endorse use of the word "priming" by Snyder. Priming in cognitive psychology has nothing to do with prime numbers). We are regrettably forced to conclude that the details of that report cannot be trusted. Considering the sincerity of Oliver Sacks in his other reports, I am inclined to believe that at least the twins really generated 6-digit primes. At the same time it is understandable that some might be skeptical about the entire report.

What can be concluded?

We do not have enough evidence to reach any definitive conclusion. However, as most of the literature on this topic contained errors, this article will redirect us to think in the right direction. One tentative conclusion is provided below.

To reiterate, although there are strong recent algorithms, imagining that the twins were the first to unconsciously derive such an algorithm would be too close to science fiction, not serious science. We must preclude these newer methods, though it is worthwhile to note that the Miller-Rabin algorithm is rather simple yet very strong.

A method potentially available to the twins was the Fermat test, which Rmachandran and Blakeslee (1998) seemed to propose. Fermat's theorem was discovered hundreds of years ago, and it is suited to primality tests for huge numbers. It can be used for such huge numbers that finding factors takes too much time even for a computer (see Appendix). Even though it is probabilistic, unless we test it extensively hundreds of times, the test will not fail. However, whether the twins found it in a book and used it is questionable. (If they could understand it, their diagnosis of impairment would have been wrong). In any event, researchers in this field are recommended to look at the Appendix and familiarize themselves with the Fermat test.

We must turn to straightforward methods. As the twins tested single numbers, it will be more reasonable to consider trial division than the sieve of Eratosthenes. As mentioned above, Sacks claimed that they seemed not to understand division. However, there are reports that impaired people show highly context-dependent abilities. For instance, Howe and Smith (1988) published a relevant case. They reported a retarded boy who was capable of calendar calculation. When posed a subtraction problem 1981−1963, he guessed 9,000 and then 3. However, when the problem was posed in the form of chronology ("If I was born in 1963, how old would I be in 1981?"), he correctly answered it most of the time. Questioned in this way, he could even correctly answer a more difficult question 2302−1841. It is very interesting that exactly the same ability was reported in a popular magazine about the twins (Hamblin, 1966), although the reaction time was far from "lightning fast". Therefore, we cannot conclude that the twins were incapable of division, but they might have had context-dependent division abilities.

As already seen, trial division requires more than 1000 operations for numbers beyond 63,000,000. Sacks reported that the twins' abilities went far beyond that range. However, details of his report are likely to be inaccurate. Therefore, we may tentatively conclude that Sacks observed the twins generated 6-digit primes, but that their abilities beyond that range are questionable. In addition, even if trial division were incomplete, their abilities might have seemed impressive. For instance, we already saw that more than 150 divisions are necessary for just under 1,000,000 for rigorous testing.

However, even if trial division is not extensive enough but consists of divisions by only the first several dozens of primes, primality can be tested relatively precisely, with only occasional errors. This is because a randomly chosen large number is more likely to be divisible by a small prime (e.g., 2 or 3) than by a large prime (e.g., 7789). Then, such abilities do not necessarily contradict our understanding of human capacities. Incomplete trial division will be a reasonable possibility, which seems consistent with the view expressed by Dehaene (2001). It is parsimonious and precludes the need for explanations by more complex probabilistic primality tests.

Finally, the discussions above are all concerned with serial algorithms. We can speculate about some parallel algorithms that can explain these surprising abilities, although of course it will not be serious science only to imply "some parallel algorithm" without specifying it.

References

Anderson M., O'Connor N., Hermelin B., (1998). A specific calculating ability. *Intelligence, 26*, 383–403

Dehaene, S. (2001). Author's Response: Is Number Sense a Patchwork? *Mind & Language, 16*, 89-100.

Hamblin, D. J. (1966). They are idiot savants: wizards of the calendar. *LIFE, 60*, 106-8.

Hermelin B., & O'Connor N. (1990). Factors and primes: A specific

numerical ability. *Psychological Medicine, 20*, 163-169.
Howe, M., and Smith, J. (1988). Calendar calculating in 'idiot savants': How do they do it? *British Journal of Psychology, 79*, 371-386.
Ramachandran, V. S., & Blakeslee, S. (1998). *Phantoms in the brain*. New York: William Morrow.
Sacks, O. (1985). *The man who mistook his wife for a hat*. London: Duckworth.
Snyder, A. (2007). Comment on priming skills of autistic twins and Yamaguchi (2007). *Journal of Autism and Developmental Disorders*.
Stewart, I. (1975). *Concepts of modern mathematics*. New York: Harmondsworth.
Welling, H. (1994). Prime number identification in idiot savants: Can they calculate them? *Journal of Autism and Developmental Disorder, 24* , 199-207.
White, P. A. (1988). The structured representation of information in long-term memory: A possible explanation for the accomplishments of "idiots savants". *New Ideas in Psychology, 6*, 3-14.
Yamaguchi, M. (2005). Comments on the Misuse of Terminology in Savant Research: It is not the Sieve of Eratosthenes. *Journal of Autism and Developmental Disorders, 35*.
Yamaguchi, M. (2007a). Questionable aspects of Oliver Sacks' report. *Journal of Autism and Developmental Disorders*.
Yamaguchi, M. (2007b). Author response. *Journal of Autism and Developmental Disorders*.

Appendix

Fermat's theorem will not seem very simple for the novices. One of the best ways to understand it will be to implement the algorithm in spreadsheet and test it with their own hands. Here the Fermat test is implemented in Excel. Of course, Excel is not academic mathematical software. Serious efforts to test primality of large numbers should use mathematical software, but for educational purposes Excel is fine.

We test whether 9991 is prime or not. First input 9991 in A1, and input 2 in C1, which is a base. As 2 must be raised to $n-1$, in A2, input =A1−1. From B1 downward, 2^n (n starts with 0) should appear. Therefore, input 1 in B1, then in B2 input =B1*2, and copy it downward, until just before 9990 is exceeded (in this case, until B14). Next in C2 input =mod(C1^2,A$1) and copy it downward. (If we naively try to compute mod(2^(2^n),9991), it quickly overflows). This process is shown in the text.

Then we must express 9990 as the sum of the numbers in column B (see the text). Although it may be done manually, here we do this automatically. This would seem technical and may be nonessential. In F14, enter =int(A$2/B14). In G14, enter =A$2−B14*F14. In F13 enter =int(G14/B13). In G13, enter =G14−B13*F13. Then copy the two cells, F13 and G13, upward.

The rest of the works may also be done manually, but here we do it automatically. In D1, enter =C1^F1, and copy it downward. If we multiply all the numbers in column D, we succeed in computing $a^{n-1} \pmod{n}$, although it quickly overflows if done naively. In E1, input =D1. And finally, in E2, enter =mod(E1*D2,A$1), and copy it downward. (See the table below).

The bottom number in the column E is the result of the Fermat test. As this is not 1, 9991 is concluded to be composite. However, the Fermat test does not tell what can divide 9991, which is actually 97*103. Considered from another angle, the Fermat test is so powerful that it can in reasonable time test primality of numbers that are too huge to test trial division to look for factors in reasonable time.

Even if the result is 1, it is not a proof of primality. We should continue testing with different bases. For instance, test 341(=11*31) with base 2, and it is erroneously judged as prime. But see that it is revealed to be composite with base 3. In principle, passing the Fermat test many times still does not prove primality, but practically the number is prime more than 99% of the time (see text).

	A	B	C	D	E	F	G
1	9991	1	2	1	1	0	0
2	9990	2	4	4	4	1	0
3		4	16	16	64	1	2
4		8	256	1	64	0	6
5		16	5590	1	64	0	6
6		32	6243	1	64	0	6
7		64	158	1	64	0	6
8		128	4982	1	64	0	6
9		256	2680	2680	1673	1	6
10		512	8862	8862	9473	1	262
11		1024	5784	5784	1188	1	774
12		2048	4788	1	1188	0	1798
13		4096	5590	1	1188	0	1798
14		8192	6243	6243	3362	1	1798