

Re: surrogate factoring

Source: <http://sci.tech--archive.net/Archive/sci.math/2006-06/msg01292.html>

- *From:* gjedwards@xxxxxxxx
 - *Date:* 8 Jun 2006 05:44:53 -0700
-

Many Thank for the extensive reply.

Tim Peters wrote:

[gjedwards@xxxxxxxx]

A previous poster was right – having anything to do with JH's posts is like rubbernecking a traffic accident. Sad but difficult to resist. Anyway ...

The bit I don't get (well one of the bits) is what specifically JH *thinks* is the point of surrogate factoring. I know it's total nonsense but does anyone have a clue *why* he thinks it's worthwhile? I guess I find the nature of the delusion somewhat puzzling and it's an amusing diversion to work out what he actually thinks.

I really should get out more.

Me too, but so long as we're both house-ridden today, and I've enjoyed more quality "SF time" than anyone other than James ...

First, "surrogate factoring" doesn't really mean anything specific. There are literally dozens of distinct methods James has _called_ "surrogate factoring" over the past two years. While they've all been wrong (in the sense that none have been efficient factoring methods, and for some it was a miracle if they ever found a factor), "total nonsense" is either overstatement or understatement :-)

The general idea is that you want to factor T, but factor some "related" integer S ("the surrogate") instead and hope to relate the factors of S to T's factors. That's perfectly sensible so far as it goes. For example, many methods try to factor $k \cdot T$ instead, and that can even be useful "by eyeball". Consider $T=817$. A tiny bit of thought shows it's not divisible by 2, 3, or 5, and then you might notice it's about a third of 2500. Indeed,

$$3 * T =$$
$$3 * 817 =$$

Re: surrogate factoring

$$\begin{aligned}2451 &= \\2500 - 49 &= \\50^2 - 7^2 &= \\(50+7)(50-7) &= \\57 \cdot 43 &\end{aligned}$$

by inspection, so you can know almost at once that 43 and $57/3 = 19$ are factors of T , if only you think about $3 \cdot T$ instead of T .

More, variations on that can be turned into straightforward algorithms that always find factors (Google for "Fermat's method"). Some of them don't even need division, which may be surprising at first sight.

Alas, trial division is usually more efficient (if you suspected I put some care into picking the 817 example, you were right ;-)), and a few of James's SF attempts have been inefficient variations on this theme, excruciatingly obfuscated ways of trying to find integers I and J such that $I^2 - J^2 = k \cdot T$, in which case $\gcd(I \pm J, T)$ has a good chance of revealing a non-trivial factor of T .

Most typical last year were SF variants that derived a large (compared to T) S to factor, and while James didn't know how to do this, it actually was possible to find S of the required form efficiently so that S could be factored efficiently despite it being much larger than T (and wrt "much larger", sometimes S grew like the 10th(!) power of T). Most criticism of those methods focused on a wrong thing, the supposed difficulty of factoring S . That wasn't the problem, but neither was it was obvious that wasn't the problem.

James insisted he had proof that given the prime factorization of S , it must be the case that some 2-integer factorization of S , $S = f_1 \cdot f_2$, revealed a factor of T after plugging f_1 and f_2 into some messy formula (which could change daily). That came in for more off-target criticism on the grounds that S may be expressible in a great many ways as the product of two integers. While that's true, it was also the case (although James didn't know how to do this either) that it was possible to find S of the required form such that S was both easy to factor and had few distinct prime divisors (the latter relates to how many ways there are to express S as the product of two integers).

The real problem was that he had no proof in reality (although he claimed to on several occasions, in his charmingly cautious way ;-)), and it was typical that no way of expressing S as $f_1 \cdot f_2$ revealed a factor of T .

Undaunted, some of us went on to search for an S of the required form such that some $f_1 \cdot f_2 = S$ existed revealing a factor of T . All such algorithms of that nature performed worse (in number of attempts needed, and extremely worse in terms of time needed given the complexity of the operations and the large size of the integers) than picking integers I at random and seeing whether $\gcd(I, T)$ revealed a non-trivial factor of T . Those weren't really James's algorithms, though — they were just enjoyable time-wasting

Re: surrogate factoring

exercises to see whether anything could be salvaged.

After that, IMO all subsequent SF attempts got worse. James got too frustrated by working with integers, and started talking about "rational factors" instead. That was silly on the face of it, since given any integer T , every non-zero rational r is "a rational factor" of T . Nevertheless, James went through increasingly elaborate piles of formulas showing how a rational factor of S could be transformed into a rational factor of T . But no transformation was necessary: every non-zero rational r is a rational factor of both S and T , for any S and T . His arguments for why this had to be a promising approach were just bizarre, and none of the resulting "methods" were actually better than this one:

Given T , pick some non-zero rational r out of the air.
If $1 < \gcd(\text{numerator}(r), T) < T$, stop, else try again.

While you could find rational factors of S to drive that, there was no point, since the set of all non-zero rationals constituted the search space regardless.

That seemed to mark the start of what's still going on, but in much slower motion this year:

1. Take a pile of starting equations as formal identities.
2. Push them around endlessly: solve for some symbols in terms of others, multiply pairs from time to time, complete the square on occasion, and sometimes invoke the quadratic formula. Don't bother restricting yourself to operations that must yield integers, or even reals — any transformation valid in the complex field is fair game.
3. Announce that "the factoring problem is solved", and the end of civilization as we know it is at hand.
4. After time passes, insist that if nobody has found factors using it yet, it's because they're too stupid to understand the simple math, but be assured that other nameless people are working on it and success is imminent. Or James is really the only person in the world smart enough to make it work, but he cares about humanity too much to risk factoring anything (but not so much that it will bother him if someone else does). The excuses here change but are always a hoot :-)
5. "Discover" that the result of #2 is really just a useless respelling of the equations you started with in #1, and the only way to make them work efficiently is to know T 's factorization before you start. This branches in two directions then:

Re: surrogate factoring

5a. Add more independent variables to the same equations, and go back to #2.

– or –

5b. Go back to #1.

The _current_ round (although James may have announced a new method or two while I was typing this) is _almost_ at step #5 in the template above. Anyone who paid attention to Rick Decker's and my postings this time knows that #5 has been solidly demonstrated to be case, but last I saw James was still fighting that. I'm not sure he's ever gotten to #5 without discovering it for himself; until then, he's always got _some_ reason to believe it's more likely that others are lying than that he's just blinded himself with hope again.

If nothing else, I hope I've convinced you that you didn't really waste your life by staying out of this :-)

As to why he does it, I think he's told us the truth on occasion: he enjoys pushing formulas around, and it may be a bona fide compulsion for him. As to why he also has a demon compelling him to believe that the results of idly pushing formulas around must be "a breakthrough", your guess may be better than mine.