



Sudoku Players' Forums

[FAQ](#) [Search](#) [Memberlist](#) [Usergroups](#)

[Profile](#) [You have no new messages](#) [Log out \[denis_berthier \]](#)

THE REAL DISTRIBUTION OF MINIMAL PUZZLES

Goto page [Previous](#) [1](#), [2](#), [3](#) ... , [12](#), [13](#), [14](#) [Next](#)



[Sudoku Players' Forums Forum Index -> General/puzzle](#)

[View previous topic](#) :: [View next topic](#)

Author

Message

Red Ed

Posted: Tue Jul 21, 2009 5:50 pm Post subject:



Joined: 06 Jun 2005
Posts: 642

Denis! How's it going? I like debating with you because I often get a quick response. Not always one I want, but a quick one nonetheless.

Re suexg & Allan's generator: which question? If it's the one about complexity stats being similar then I already answered: I don't know (or care). Otherwise, do please remind me.

He he he, btw: I like your $2+2=4$ or 5 thing 😊 But okay, let's be serious. Do you honestly not believe the statistical [evidence](#) that I presented? What's your objection to it? What more would convince you? Or will you concede the point?

denis_berthier wrote:

Red Ed wrote:

after all, we already have strong evidence that the *suexg* solution grid generator produces biased number-of-minimals estimates

Obviously, we don't have the same notion of "evidence". The only evidence we have is that the *suexg* generator of minimal puzzles is biased - that's not a scoop. We have no evidence that its complete grids generator part is biased.

[Back to top](#)



denis_berthier

Posted: Tue Jul 21, 2009 6:41 pm Post subject:



Joined: 19 Jun 2007
Posts: 770
Location: Paris, France

Red Ed wrote:

Re *suexg* & Allan's generator: which question? If it's the one about complexity stats being similar then I already answered: I don't know (or care).

You should care because they show that your objections about bias in the complete grids are irrelevant wrt complexity - in the same way as your 3322

tests or so were irrelevant.

As for complete grids, as I told you several times, *I don't care*. At least two thirds of the grid are deleted before a minimal puzzle is obtained.

I'm very confident that the controlled-bias generator would produce exactly the same results if we changed its complete grids part (keeping it sufficiently random, of course).

[Back to top](#)

[profile](#) [pm](#) [www](#)

Red Ed

Posted: Tue Jul 21, 2009 7:18 pm Post subject:

[quote](#)

Joined: 06 Jun 2005
Posts: 642

Ah-ha! Denis, good, good, you've answered my question:

denis_berthier wrote:

I'm very confident that the controlled-bias generator would produce exactly the same results if we changed its complete grids part (keeping it sufficiently random, of course).

"Exactly the same" - a bold claim! 😊 And one that I am sure I can prove wrong.

Let's agree upon an experiment. I'll modify suexg-cb.c to take an external stream of grids as input, then I'll run it on a couple of different sources (my generator plus at least one of: { Allan's puzzles, the sudogen0_1M solution grids, the suexg-cb.c built-in generator }). I'll do a distribution equality test (e.g. KS test, and/or another of your choosing) and will, I think, be able to prove to 95% confidence level or better that the results are different.

I see this as a test of credibility. You are "very confident" about suexg-cb's lack of dependence on its solution grid source. Do you wish to retract that, or go through with the experiment?

[Back to top](#)

[profile](#) [pm](#)

denis_berthier

Posted: Tue Jul 21, 2009 10:35 pm Post subject:

[quote](#) [edit](#)

Joined: 19 Jun 2007
Posts: 770
Location: Paris, France

Red Ed,

I can understand you need to pass a test of credibility. After working so long on the number of clues distribution problem, you missed completely the very simple idea of a controlled-bias generator.

So let me define the conditions of the test. The general idea is: instead of blah blah-ing, work.

1) Forget about Allan's generator. I've already proven, much before our discussion started, that it gives the same results (mean number of clues, mean SER, mean NRCZT) as the standard top-down suexg. There's no reason that doing the modification for making it controlled-bias would change much of this.

2) Modify the complete grids part of suexg-cb by introducing the code for your supposedly unbiased generator. Make it public (and give a chance to your

repeated claims that you want to be helpful) or get lost. I've enough of your unsustained claims that no one can check.

3) When done, generate 10,000 puzzles with what will still be a controlled-bias generator. Make them public or get lost.

4) When done, run your KS test.

[Back to top](#)

 [profile](#)  [pm](#)  [www](#)

Red Ed

📄 Posted: Wed Jul 22, 2009 5:46 am Post subject:

 [quote](#)

Joined: 06 Jun 2005
Posts: 642

denis_berthier wrote:

Make them public or get lost.

He he he, you are an angry little man aren't you? 😊

No, Denis, I shan't make my generator public yet, as I have received no interest in it except from you ... and you've been so *rude*. I suppose you must be frustrated that you don't understand what's going on around you. It's a confusing world, isn't it, Denis? There, there. Have a lollipop.

I wonder what your students think of this. They should be rather embarrassed that you've invested so much CPU time in:

- A method that's biased (see evidence quoted before; I can explain it to you if you find it difficult).
- Answers quoted to 5 sig fig without any estimate of the variance, only a weak warning to the reader to exercise caution.

Not very scientific, Denis. Maybe you should ask your students for help. Someone in your establishment must know some statistical methods, surely.

Other readers on the Players' Forum must be quite bored of us by now. Maybe we should stop this lovers' tiff, darling.

[Back to top](#)

 [profile](#)  [pm](#)

denis_berthier

📄 Posted: Wed Jul 22, 2009 6:03 am Post subject:

 [quote](#)  [edit](#)

Joined: 19 Jun 2007
Posts: 770
Location: Paris, France

Red Ed,

Always the same insults and evasive answers when it comes to justifying your claims or your opaque methods.

So little imagination is boring.

In fine, your method is very simple:

- claim that your generator of complete grids is unbiased,
- compute various statistics with it,
- use these statistics to prove that your generator is unbiased,
- use these statistics to prove that everybody else is wrong.

You're indeed a master in statistics and scientific method. Teach me, darling!

[Back to top](#)[profile](#) [pm](#) [www](#)**Red Ed**

Posted: Wed Jul 22, 2009 6:04 am Post subject:

[quote](#)Joined: 06 Jun 2005
Posts: 642

Now, diatribe aside, is there any technical point on which we disagree apart from the sensitivity of the controlled-bias top-down generator to its source of solution grids? I think it's sensitive (affecting the number-of-clues distribution by a few percent), you think it's not ("exactly the same results").

Is that really all we're bickering about?

[Back to top](#)[profile](#) [pm](#)**denis_berthier**

Posted: Wed Jul 22, 2009 6:27 am Post subject:

[quote](#) [edit](#)Joined: 19 Jun 2007
Posts: 770
Location: Paris, France**Red Ed wrote:**

Now, diatribe aside, is there any technical point on which we disagree apart from the sensitivity of the controlled-bias top-down generator to its source of solution grids? I think it's sensitive, you think it's not.

Oh, so you're finally accepting the idea of controlled-bias? What a great step forward for mankind! 😊

Sensitivity is of course a point of disagreement. But notice that, even if it was sensitive (which is very unlikely), that would in no way change the idea I had of a controlled-bias generator. We'd just have to improve the source of solution grids. Fusing your supposedly unbiased complete grid generator with my idea of a controlled-bias may have been a way of doing this (at least, it should have stopped your recriminations). But I've well understood from your previous declaration of love that you don't want to participate in this. You love nitpicking and bickering but no one can rely on you for constructive participation. That's a second point of disagreement.

But there is also a point of method: whereas I've always made available (through my website) all the elements that made it possible to check my results (and you largely relied on this in your criticisms), you're playing an opaqueness game that's really boring.

I don't know why, but I love talking with you, anyway. I must be very masochistic. 😞

[Back to top](#)[profile](#) [pm](#) [www](#)**Red Ed**

Posted: Wed Jul 22, 2009 6:34 am Post subject:

[quote](#)Joined: 06 Jun 2005
Posts: 642

I'll ignore the personal comments, as it's clear to both of us that neither particularly likes the other's methods.

I'll find it harder to ignore the great step forward for mankind comment, although its timing (moon landings etc.) is apt. I've never said that the (new) controlled-bias generator was anything other than unbiased in principle. I just claim that it's

biased in practice owing to the solution grid source.

I will now focus on proving to you that the number-of-clues distribution is different (not by much, but different nonetheless) when the solution grid source is different. I might do that on another thread.

[Back to top](#)

 [profile](#)  [pm](#)

denis_berthier

Posted: Wed Jul 22, 2009 6:49 am Post subject:

 [quote](#)  [edit](#)

Joined: 19 Jun 2007
Posts: 770
Location: Paris, France

Red Ed wrote:

I've never said that the (new) controlled-bias generator was anything other than unbiased in principle.

One more of your word games.

We know that the puzzles produced by this generator are biased (and we know how to correct this bias). How can you call it "unbiased", even "in principle"?

Red Ed wrote:

I just claim that it's biased in practice owing to the solution grid source.
I will now focus on proving to you that the number-of-clues distribution is different (not by much, but different nonetheless) when the solution grid source is different. I might do that on another thread.

No reason to do that elsewhere, as this discussion started here, but if you prefer so, I don't mind. Just tell me where when you have finished: I'm not watching many threads.

But the question remains: which part of your forthcoming claims shall we be able to check independently?

Red Ed wrote:

the number-of-clues distribution is different (not by much, but different nonetheless) when the solution grid source is different.

I appreciate the "not by much" as an anticipated retraction of your previous claims. What about sampling error?

[Back to top](#)

 [profile](#)  [pm](#)  [www](#)

Red Ed

Posted: Wed Jul 22, 2009 7:14 am Post subject:

 [quote](#)

Joined: 06 Jun 2005
Posts: 642

denis_berthier wrote:

Red Ed wrote:

I've never said that the (new) controlled-bias generator was anything other than unbiased in principle.

One more of your word games.

We know that the puzzles produced by this generator are biased (and we know how to correct this bias). How can you call it "unbiased", even

"in principle"?

Not a word game, Denis; just a mistake. I meant known bias, not absence of bias; so that unbiased *estimators* can be constructed from it. Please don't whinge: I was giving it the thumbs-up.

As the for rest, stop clogging up the forum with waffle and just wait and see please.

[Back to top](#)

[profile](#) [pm](#)

denis_berthier

Posted: Wed Jul 22, 2009 7:28 am Post subject:

[quote](#) [edit](#)

Red Ed,

Joined: 19 Jun 2007
Posts: 770
Location: Paris, France

Thanks for teaching me a lot of offensive English vocabulary.

[Back to top](#)

[profile](#) [pm](#) [www](#)

Red Ed

Posted: Wed Jul 22, 2009 1:47 pm Post subject:

[quote](#)

Joined: 06 Jun 2005
Posts: 642

No problem. Probably no point us doing the same in French as I'm as bad at that as you say you are at 'C'. *Tant pis.*

[Back to top](#)

[profile](#) [pm](#)

denis_berthier

Posted: Sat Jul 25, 2009 11:35 am Post subject:

[quote](#) [edit](#)

denis_berthier wrote:

Joined: 19 Jun 2007
Posts: 770
Location: Paris, France

I'm very confident that the controlled-bias generator would produce exactly the same results if we changed its complete grids part (keeping it sufficiently random, of course).

This prediction about the controlled-bias generator is motivated by two reasons:

- two thirds of a complete grid are eliminated during the deletion phase;
- more than 100,000 complete grids are necessary in the mean to have the controlled-bias generator produce one minimal puzzle; any bias that might have existed in the source of complete grids is thus very likely to be eliminated by this randomisation process.

As Red Ed started out challenging this prediction in another thread:

Red Ed wrote:

I will now focus on proving to you that the number-of-clues distribution is different (not by much, but different nonetheless) when the solution grid source is different. I might do that on another thread.

it seems appropriate, for completeness of the present thread, to recall here the conclusions we reached there:

Red Ed wrote:**denis_berthier wrote:**

Said otherwise, any "sufficiently random" complete grids generator (e.g. that of suexg) will do as well (for puzzle generation) as a more sophisticated one.

Oh, sure. It's become pretty clear that the bias in the proportion-of-n-clue-minimals distribution is very small. No-one is going to notice in practice unless they're making a deliberate effort to go looking.

We could even conclude to insensitivity wrt a large bias in the complete grids:

Red Ed wrote:**denis_berthier wrote:**

You said A is your "unbiased generator" and B is suexg's. So suexg would produce complete grids with 20% more minimal than normal, right?

Yep, the bias is real.

Red Ed wrote:**denis_berthier wrote:**

even a strong (20%) bias in the mean number of minimal of a stream of complete grids used as input to a controlled-bias generator of minimal puzzles has only a very small impact on the distribution of clues of the output.

Certainly appears to be true for the *suexg* input.

It seems all this gives a clear confirmation to the above prediction.

[Back to top](#)
 [profile](#)
 [pm](#)
 [www](#)
Red Ed

Posted: Sat Jul 25, 2009 1:47 pm Post subject:

 [quote](#)

Joined: 06 Jun 2005
Posts: 642

Mmm... not so fast. You wrote "exactly the same". The distributions are not exactly the same. I concur that they're close enough that almost no-one should ever care about the difference, though.

[Back to top](#)
 [profile](#)
 [pm](#)

Display posts from previous:

 [newtopic](#)
 [postreply](#)

Sudoku Players'
Forums Forum
Index ->
General/puzzle

All times are GMT
Goto page [Previous](#) [1](#), [2](#), [3](#) ... , [12](#), [13](#), [14](#) [Next](#)

Page 13 of 14

[Stop watching this topic](#)

Jump to:

You **can** post new topics in this forum
You **can** reply to topics in this forum
You **can** edit your posts in this forum
You **can** delete your posts in this forum
You **can** vote in polls in this forum

Powered by phpBB © 2001, 2005 phpBB Group