Rating rules / Puzzles. Ordering the rules

Goto page Previous 1, 2, 3 ... , 24, **25**, 26 Next

newtopic postreply **Sudoku Players' Forums Forum Index -> Advanced solving techniques**

View previous topic :: View next topic

Author | **Message**

StrmCkr

Posted: Thu Jul 09, 2009 5:51 pm Post subject: quote

Joined: 05 Sep 2006
Posts: 491
Location: Winterpeg

Quote:

> Now you say, when you arrive at the first number, which cannot be eliminated, you replace it by another one, thus building local minimal in 2 grids now. And if you take the lower one, this would "produce the results of a bottom up generator".
>
> Why only the first number, why [edit] the number you cannot drop at this stage ?
> And what does that have to do with bottom up generation ?

in many grids there is several or few locals that can represent a completed grid,

{I've also shown this in the stormduko variation thread }

for example.
11.4 #
500000009020100070008000300040600000000500000002070100030008000600040209000000005
m_b_metcalf
11.4 #
500000009020100070008000300040002000000050000000706010003000800060004020900000005
StrmCkr

both grids are minimal, i found mine by deleting and adding 1 clue from m_B_metcalfs; and generated this none isomorphic puzzle.

the difference in a top down generator is that the puzzle produced by the generator all hold the same type of constructs in a grid based on the logic assigned from the start.

with that any singular clue removed or deleted confirms that the cover set still leaves 1 solution grid.

with +1-N,-1-N you can modify the complexity of the logic sets and as a by product the minimal clues needed to satisfy 1 solution can be lowered.

the end results often is a easy puzzle.

which is what happens in a bottom up generator it builds the constructs of logic and may destroy any complexities it was randomly building by placing a key digit on the grid.

Quote:

> Or maybe you're trying to generate minimal puzzles in an unbiased way. In which case I'll repeat what I said on the "How many minimal sudokus has an average grid" thread

> yesterday, that the average number of clues that you should be aiming for is in the region of 26.something. You're a long way off that ...

red ed, im curious if the random sample cannot represent the average in whole due to the shear amount of different grids one can produce.

with each sample you should generate different representation of samples thus slightly different statistics.

the larger % comparison to the real total number of grids the more accurate the mean of minimal would be.

1 million is fractions of the real number.

6,670,903,752,021,072,936,960

Back to top profile pm

Allan Barker

Joined: 20 Feb 2008
Posts: 296
Location: Bangkok

Posted: Thu Jul 09, 2009 7:47 pm Post subject: quote

eleven wrote:

> Also only the mean clue value really compares to the bottom up distribution. E.g. in your sample there are 5% more 24's (41.6 to 36.4).

Yes, your correct, the distribution is tighter overall, which is interesting. I edited my post to reflect that I'm talking about mean average clue sizes.

eleven wrote:

> I dont understand this. [.....] what does that have to do with bottom up generation ?

Although all steps are random, including clue sizes, the argument is that for a given clue size the top down has fewer chances (than the bottom up) to find a minimal puzzle, exactly 1 less in this case because A's outcome is known. Experiment 3 gives the top down an equal opportunity (to the bottom up) to find a local minimal *without disturbing* the rest of the process. At which point, it produces puzzles with the same average clue size (as the bottom up).

So, I think the difference is just sampling opportunities, otherwise this difference wouldn't always come out so close to ½. But as **Red Ed** mentioned the subject seems to be drifting. Is there a better place ?

Back to top profile pm www

denis_berthier

Joined: 19 Jun 2007
Posts: 760
Location: Paris, France

Posted: Thu Jul 09, 2009 9:09 pm Post subject: quote

Red Ed wrote:

> Are you all just trying to understand why different puzzle generators exhibit different statistics? Or trying to generate puzzles with few clues? Both perfectly reasonable aims, but surely unrelated to the original topic.

Trying to generate puzzles with few (or many) clues is clearly off topic.

Trying to understand why different random generators give (slightly) different statistics is not off topic if it helps understand whether there is a bias in generators wrt to the difficulty of puzzles and why there is one. I think there are now a few things we understand better.
It can also show that, in spite of possibly different biases, some results remain true for all the generators, e.g. the strong correlation between the SER and NRCZT ratings, in spite of their being based on completely different sets of rules.

Red Ed wrote:

> Or maybe you're trying to generate minimal puzzles in an unbiased way.

That's an ultimate goal - but just saying "unbiased" is meaningless. We only need them to be unbiased wrt the complexity of puzzles. So that, e.g., the number of occurrences of any of the patterns in your list of 3322 is not relevant (as it is uncorrelated to complexity).
The mean number of clues and the distribution of this number is not irrelevant, because it seems to have a (weak) correlation with complexity. It is marginally relevant.
In this respect, the correlation coefficient, a very elementary tool, can easily eliminate from consideration many irrelevant biases.

Red Ed wrote:

> In which case I'll repeat what I said on the "How many minimal sudokus has an average grid" thread yesterday, that the average number of clues that you should be aiming for is in the region of 26.*something*. You're a long way off that ...

You weren't so affirmative there:

Red Ed wrote:

> That's a waaay higher average number of clues than what we're used to estimating. I think my method's unbiased (will rethink to be sure), so maybe it just needs longer to converge. Or maybe the average number of clues in a minimal puzzle really is that high.
>
> EDIT: I'm also suspicious of the "Total minimals per grid" figure, since it's less than my 03-Dec-2006 estimate of 6.6e15 minimals by a method that I think is equivalent to Knuth's 1975 random probing algorithm -- which is known to give underestimates. So take all this with a pinch of salt.

With only 0.49% of your sample above the 26.4 mean, you were certainly right to show some caution.

Allan Barker wrote:

> But as **Red Ed** mentioned the subject seems to be drifting. Is there a better place ?

If discussions on puzzle generation which have no clear relation with rating or classification should continue, I agree with Red Ed and Allan that they should be in another thread, so that we can concentrate here on these topics.

Back to top profile pm www

Red Ed

Posted: Thu Jul 09, 2009 10:42 pm Post subject: quote

Joined: 06 Jun 2005
Posts: 634

All fine what you said above except that you couldn't resist a "meaningless" barb:

denis_berthier wrote:

> **Red Ed wrote:**
>
> > Or maybe you're trying to generate minimal puzzles in an unbiased way.
>
> That's an ultimate goal - but just saying "unbiased" is meaningless.

It is not meaningless, Denis, as you well know. "Unbiased" as in "samples uniformly at random from the population", if you need me to spell it out. If you had such a generator at your fingertips, and it was tolerably quick, then we both know you'd be using it.

Back to top profile pm

denis_berthier

Posted: Thu Jul 09, 2009 11:07 pm Post subject: quote

Red Ed wrote:

Joined: 19 Jun 2007
Posts: 760
Location: Paris, France

All fine what you said above except that you couldn't resist a "meaningless" barb:

denis_berthier wrote:

> **Red Ed wrote:**
>
> > Or maybe you're trying to generate minimal puzzles in an unbiased way.
>
> That's an ultimate goal - but just saying "unbiased" is meaningless.

It is not meaningless, Denis, as you well know. "Unbiased" as in "samples uniformly at random from the population", if you need me to spell it out. If you had such a generator at your fingertips, and it was tolerably quick, then we both know you'd be using it.

No "barb" intended.
Of course, I (and anyone interested in statistics) would use such a generator.
As the rest of the sentence suggests, I only meant meaningless in practice (because unreachable) for the purposes of this thread. We must concentrate on reachable goals.

Back to top profile pm www

denis_berthier Posted: Fri Jul 10, 2009 2:30 am Post subject: quote

Joined: 19 Jun 2007
Posts: 760
Location: Paris, France

denis_berthier wrote:

> The previous computations have shown that puzzles generated bottom-up tend to be easier than puzzles generated top-down.
> ...
> I imagined the following double generation process:
> 1) take the first part of a bottom-up generator and get a non complete grid G1 with a unique solution
> 2) with the second part of the generator obtain a minimal puzzle P1
> 3) let G2 be the complete (solution) grid obtained from G1
> 4) with exactly the same second part (this is very important) obtain a minimal puzzle P2
>
> Question: is there any correlation between P1 and P2:
> - number of clues
> - SER
> - NRCZT

m_b_metcalf wrote:

> I added just three statements to my program, and got the following result in 500s:
>
> **Code:**
>
> ```
> Bottom-up
> ---------
> Number:        11177
> Average:    23.88798
>
>      20        0
>      21       94
>      22      918
>      23     3041
>      24     4040
>      25     2362
>      26      631
>      27       81
>      28       10
>      29        0
>
> Top-down
> --------
>
> Number:        11177
> Average:    24.37049
>
>      19        0
>      20        1
> ```

```
21      28
22     390
23    1985
24    3779
25    3332
26    1354
27     279
28      28
29       1
30       0
```

denis_berthier wrote:

> That's really very interesting. You've definitely proven that the difference for the number of clues is not related to the specifics of the deletion phase.

This was before we got general explanations for this top-down vs bottom-up diffference in the number of clues.
The computations (on 2x10,000 puzzles) for the SER are now over.
As expected, we get an almost null P1 P2 correlation for the numbers of clues and a null correlation for the SER:
#clues(P1) vs #clues(P2) = 0.027
SER(P1) vs SER(P2) = 0.0007

How can one interpret this?
Given a bottom-up minimal P1 and its solution G1, the deletion phase from G1 to the minimal P2 erases almost any structure that came from P1.
Well, this is not really a scoop.

Back to top profile pm www

denis_berthier

Posted: Mon Jul 13, 2009 10:47 pm Post subject: quote

Joined: 19 Jun 2007
Posts: 760
Location: Paris, France

Back to the main topic of this thread, classification.

After I reported detailed statistical results for the sudogen0_1M collection of 1,000,000 minimal puzzles generated by suexg version x.x (here http://www.sudoku.com/boards/viewtopic.php?t=5995&postdays=0&postorder=asc&start=142), there has been some discussion about other generators.
I decided to check whether my results could be extended to large collections produced by different generators, 2 top-down (suexgx.x and one by Allan Barker) and 2 bottom-up (suexg14 and one by Mike Metcalf).

I reported preliminary results here: http://www.sudoku.com/boards/viewtopic.php?t=5995&postdays=0&postorder=asc&start=318

Most of the computations that have been running since then in the background (when the CPU was not busy for anything else) are now finished.

The SER ratings have been computed for the complete 4 collections.
The NRCZT ratings have been computed for larger sub-collections than in the "preliminary results" (e.g. 50000 instead of 10000 for Allan's generator).

These extended computations only confirm the conclusions I gave in the "preliminary results".

As I can't put the detailed results in html tables here, I've just updated my web page with them (http://www.carva.org/denis.berthier/HLS/Classification) . There, you can find all the tables of results and you can access all the files justifying them (for all the collections: SER, NRCZT, #clues, ...).

Let me just recall here the main conclusions:

*** the close similarity between the 2 top-down generators:**
o same mean number of clues,
o same distribution of the number of clues,
o close mean SER and mean NRCZT, globally and for each number of clues;
*** the close similarity between the 2 bottom-up generators:**
o same mean number of clues,
o same distribution of the number of clues,
o close mean SER and mean NRCZT, globally and for each number of clues;
*** the clear difference between the top-down and bottom-up generators:**
o notably larger mean number of clues for the top-down generators (+ 1/2),
o notably different distributions of the number of clues,
o globally larger mean SER and mean NRCZT for the top-down generators;
*** for the 4 generators, a number of clues between 20 and 30 for all the puzzles** (exceptionally, one puzzle with 19 clues with Allan's generator);
*** for the 4 generators, a small trend for increasing SER or NRCZT with increasing number of clues between 21 and 29** (not enough data to allow any conclusion below 21 or above 29 - except for sudogen0_1M, for which the trend appears in the whole 20-30 span);
*** nevertheless, for the 4 generators, a very small (~0.1) correlation coefficient between the number of clues and the SER or NRCZT, which implies that the number of clues can't be used for predictions of the (SER or NRCZT) complexity of an individual puzzle**.

The following general results are true for all the generators considered above:

*** all the minimal puzzles built with random generators can be solved by nrczt-whips;**
*** at least 99% of these minimal puzzles can be solved by nrczt-whips of length 5 or less;**
* **at least 99,9% of these minimal puzzles can be solved by nrczt-whips of length 7 or less;**
*** the SER rating provides (after conversion) a *statistically* good approximation (correlation coefficient ~0.895) of the (harder to compute, but purely logic and intrinsic) NRCZT-rating.**

Back to top profile pm www

Red Ed

Joined: 06 Jun 2005
Posts: 634

Posted: Wed Jul 22, 2009 9:45 am Post subject: quote

David P Bird wrote:

> FWIW here's a puzzle grid generation scheme I have manually tried several times now without reaching any invalid grids - mind you, I don't run at 2GHz and for random selection read "mental coin toss" so I can't really claim it's infallible.
>
> 1) Randomly fill boxes 1 & 5 with digits 1 to 9
> 2) Randomly select digits to occupy r7c7, r8c8 and r9c9.
>
> From this point on, simple Sudoku eliminations must be made as each new digit is placed to reduce the cells sets available for further random placements.
> The general aim is to take the digits one by one and position them randomly in the cells open to them.
> The order in which the open cells are selected for the current digit is determined by prioritising the houses with most restricted choice first.
>
> 3) For the digit in r7c7 randomly select one of the available cells open for it in boxes 2,3,4, & 7 forcing its position in boxes 6 and 8
> 4) Repeat for digit in r8c8 in a box order decided by restricted choice
> 5) Repeat for digit in r9c9
>
> 6) Now same restricted choice principle is used to determine the order in which the remaining digits are placed, and where there is a choice, it is decided randomly.
> Priority 1: Any digit which has already been forced as a single as a result of the previous placements
> Priority 2: Any digit which is bi-local in any house.

> Priority 3: One of the digits in any bivalue cell.
> Priority 4: One of the digits in the first unresolved cell in box 2
>
> Braid analysis logic lies at the back of this scheme which tells us that the same travelling pair can't exist both in a tier and a stack. The selection order used simply avoids such possibilities. Should you try it out, watch out for the emergence of unconditional sets towards the end.
>
> I can't see that this system is biased, but then I know I am, so I wait for the opinions of others.

Do you have code? I'd be happy to run some tests.

EDIT: I should've said 'C' code, or output; either's fine.

Back to top profile pm

David P Bird

Posted: Wed Jul 22, 2009 10:08 am Post subject: quote

Joined: 16 Sep 2008
Posts: 74
Location: Middle England

> **Quote:**
>
> Do you have code? I'd be happy to run some tests.

As I hoped to infer from my opening sentence, I have only used this algorithm manually. No code has been written, but for someone with the right code repository it shouldn't take too long to develop.

The question I'd like answered is whether or not this approach would satisfy your no-bias criteria, provided of course my no-backtracking needed claim were to be substantiated.

Back to top profile pm

Red Ed

Posted: Wed Jul 22, 2009 10:23 am Post subject: quote

Oh. So I need to think, not just hit "Go". :(

Joined: 06 Jun 2005
Posts: 634

I'll commit only to "very probably biased". Almost everything is, although for most applications the bias isn't noticeable. Bias only matters for stuff like estimating the number of minimal puzzles.

Back to top profile pm

gsf

Posted: Wed Jul 22, 2009 7:34 pm Post subject: quote

Joined: 21 Sep 2005
Posts: 3583
Location: NJ USA

> **Red Ed wrote:**
>
> Do you have code? I'd be happy to run some tests.

you probably did this for my solver a while ago
it uses rookeries to generate solution grids
if there are some new tests then you can try them on

> **Code:**
>
> ```
> sudoku -gg -nN > grids.dat
> ```

where *N* is the number of grids to generate
rate is about 10,000 grids/sec/Ghz

Back to top profile pm www

denis_berthier

Posted: Wed Jul 22, 2009 11:22 pm Post subject: quote

Joined: 19 Jun 2007
Posts: 760

UNBIASED CLASSIFICATION RESULTS

Location: Paris, France

Since my last post in this thread (http://www.sudoku.com/boards/viewtopic.php?t=5995&start=366), where I gave raw classification results for 4 collections of puzzles, 2 generated by top-down generators and 2 generated by bottom-up generators, much has been going on.
Let me state briefly where we now are.

It is now clear that there is a bias in favour of puzzles with few clues, both in top-down and bottom-up generators and that it is stronger in the latter. In the sequel, I'll therefore concentrate on top-down generators.
Notice that the bias explains why it is so difficult to find minimal puzzles with 30 or more clues.

This bias is known to have limited influence on the (SER or NRCZT) complexity of puzzles, because I've already shown that the correlation between the number of clues and the complexity is small (~ 1.1).

But, if you're a purist, you'd like to know more precisely the influence of this bias on the above classification results. That's the purpose of this post.

Unfortunately, it is also clear that the number-of-clues bias is very hard to analyse in the current generators.
I therefore devised a new generator, **the "controlled-bias" generator**, which is a modified version of the classical top-down generators.
Details on this generator are given here: http://www.sudoku.com/boards/viewtopic.php?t=14615&start=134; the most up-to-date version is on my web page:
http://www.carva.org/denis.berthier/HLS/Classification

Described shortly, the algorithm is as follows:

Code:

```
1) generate a random complete grid
2) loop:
     let P be the current puzzle
     2a) choose one clue randomly from P and delete it, you get a
puzzle P2
     2b) if P2 is minimal, return P2
     2c) if P2 has several solutions, GOTO 1
     2d) otherwise, set P=P2
end loop
```

It differs from the classical top-down algorithm only by clause 2c (instead of "GOTO 2a", we have a "GOTO 1"): in case the search for a minimal finds a multi-solution puzzle, the current path is merely discarded, instead of backtracking. It implies that many of the complete grids generated in phase 1 will lead to the generation of no minimal puzzle.

Eleven could easily modify the top-down version of suexg to implement this new algorithm, suexg-cb (also available on my web page). Thanks to him.

Let it be clear that this is not an unbiased generator of minimal puzzles (IMO, an unbiased generator is an unrealistic goal). It is a generator with known bias (whence the name controlled-bias), i.e. it is amenable to a precise mathematical analysis and correction factors can be defined in order to get unbiased statistics from raw statistics computed from the output of this generator (see the above mentioned references for details).

Notice that this new generator is very slow; but speed isn't the purpose here. I could nevertheless generate 13,000 puzzles (current number) - which is largely enough for the rough estimation of the number-of-clues distribution we need for our classification purposes.

Let me now report the results relevant to this thread, based on the above 13,000 minimals.

For the mean number of clues:

- bottom-up generators (suexg14): 23.87
- top-down generators (suexgx.x): 24.38
- controlled-bias generator (suexg-cb): 25.659
- unbiased (suexg-cb with correction factors): **26.577**

For the SER:
- bottom-up generators (suexg14): 3.50
- top-down generators (suexgx.x): 3.77
- controlled-bias generator (suexg-cb): 4.151
- unbiased (suexg-cb with correction factors): **4.488**

For the NRCZT:
- bottom-up generators (suexg14): 1.80
- top-down generators (suexgx.x): 1.94
- controlled-bias generator (suexg-cb): 2.137
- unbiased (suexg-cb with correction factors): **2.312**

Sensitivity analysis.
The main limitation in the precision of the number-of-clues distribution is due to the difficulty of finding minimal puzzles with 30 or more clues.
In my current sample of 13,000 minimal puzzles, there are only 4 30-clue and no c-clue for c>30.
But a sensitivity analysis is still possible.

- if, to the same sample we added a 31-clue, a 32-clue, a 33-clue, a 34-clue and a 35-clue - the worst, very unlikely case one could imagine -, the unbiased mean number of clues would be: 26.73 (instead of 26.577)

- if, to the same sample we added a 31-clue, a 32-clue and a 33-clue (no 34-clue or 35-clue added, because I don't know the mean SER for these - but still a very unlikely case), the unbiased mean SER would be: 4.506 (instead of 4.488).

Conclusion: the unbiased complexity results provided by the controlled-bias generator are very stable.

Back to top profile pm www

Red Ed

Joined: 06 Jun 2005
Posts: 634

Posted: Thu Jul 23, 2009 6:28 am Post subject: quote

gsf wrote:

> **Red Ed wrote:**
>
> > Do you have code? I'd be happy to run some tests.
>
> you probably did this for my solver a while ago
> it uses rookeries to generate solution grids
> if there are some new tests then you can try them on
>
> **Code:**
>
> ```
> sudoku -gg -nN > grids.dat
> ```
>
> where *N* is the number of grids to generate
> rate is about 10,000 grids/sec/Ghz

Very impressive: Z_1M score (bias) of 60.95, much better than *suexg*'s Z_1M score of 200+. See the unbiased grid generation thread for an explanation of that score.

Back to top profile pm

eleven

Posted: Thu Jul 23, 2009 11:23 am Post subject: quote

Joined: 10 Feb 2008
Posts: 371

Red Ed wrote:

> Very impressive: Z_1M score (bias) of 60.95, much better than *suexg*'s Z_1M score of 200+.

Can you tell me, how much grids i should have (generated with gsf's or an "unbiased" program), that i could use suexg with a "harmless" bias ?

Back to top profile pm

denis_berthier Posted: Thu Jul 23, 2009 10:15 pm Post subject: quote

Joined: 19 Jun 2007
Posts: 760
Location: Paris, France

Red Ed wrote:

> Very impressive: Z_1M score (bias) of 60.95, much better than *suexg*'s Z_1M score of 200+.

Don't forget that the Z_1M score is relevant for complete grids but not for complexity of puzzles.

eleven wrote:

> Can you tell me, how much grids i should have (generated with gsf's or an "unbiased" program), that i could use suexg with a "harmless" bias ?

- "harmless" for which purpose?
- which version of suexg are you speaking of (bottom-up, top-down or controlled-bias)? You can understand from my previous post that it makes much difference.

Considering the topics of this thread and supposing that you want to use gsf's (or any other) generator of complete grids as an input to the controlled-bias generator (instead of suexg's own complete grids generator), my (very rough) estimate (based on computation times) is that a controlled-bias minimal puzzle is worth between 2000 to 3000 ordinary top-down minimal puzzles (and it consumes the same amount of different complete grids).

You can get a better estimate of this figure by doing the following modification in suexg-cb:
- for each minimal puzzle reached by the algorithm, output (e.g. on the same line, separated by a space) the puzzle and the number of complete grids necessary to obtain it,
- at the end, it's easy to make an average.
(Don't output only the global number of complete grids at the end, because, as generation is very slow, you may want to stop it before you've got the initially desired number of minimal puzzles).

If you do this modification, in the future I'll use it so that you can take advantage of my computations for your own purposes.

Back to top profile pm www

Display posts from previous: All Posts Oldest First Go

newtopic postreply **Sudoku Players' Forums Forum Index -> Advanced solving techniques**

All times are GMT - 8 Hours
Goto page Previous 1, 2, 3 ... , 24, 25, 26 Next

Page 25 of 26

Jump to: Advanced solving techniques Go