Sudoku Players' Forums

Rating rules / Puzzles. Ordering the rules

Goto page Previous 1, 2, 3 ... , 13, 14, 15 Next

new topic post reply Sudoku Players' Forums Forum Index -> Advanced solving techniques

View previous topic :: View next topic

Author | **Message**

denis_berthier

Posted: Fri Jul 03, 2009 7:53 am Post subject: quote

Yes, but why 3322 patterns?
I guess they are not completely arbitrary.
Do they form some complete set of patterns wrt to some criterion?

Joined: 19 Jun 2007
Posts: 631
Location: Paris, France

Back to top profile pm www

m_b_metcalf

Posted: Fri Jul 03, 2009 8:06 am Post subject: quote

Joined: 15 May 2006
Posts: 2147
Location: Berlin

m_b_metcalf wrote:

> **denis_berthier wrote:**
>
> > **Mike**, can't you do something like this with your generator? If you can produce such a sample, I'd like to test it.
>
> I have a different old generator which I think I can more easily adapt to do this. But next week.

It started raining, so I made a first stab at this. I need to make some more checks and improve the speed, but a first result, if sustained, is tantalizing:

Code:

```
Number:         1760
Average:    23.87329

      20        0
      21       16
      22      141
      23      482
      24      659
      25      348
      26      103
```

```
27      9
28      2
29      0
```

and includes an SE=9.1.

Regards,

Mike Metcalf

Back to top profile pm

Red Ed

Posted: Fri Jul 03, 2009 8:59 am Post subject: quote

Joined: 06 Jun 2005
Posts: 522

denis_berthier wrote:

> Yes, but why 3322 patterns?
> I guess they are not completely arbitrary.
> Do they form some complete set of patterns wrt to some criterion?

The historical records don't go back that far ... IIRC, it was just a collection of patterns that were good at defeating the generators of the day.

Back to top profile pm

denis_berthier

Posted: Fri Jul 03, 2009 9:00 am Post subject: quote

Joined: 19 Jun 2007
Posts: 631
Location: Paris, France

m_b_metcalf wrote:

> a first result, if sustained, is tantalizing
> ...
> and includes an SE=9.1.

Interesting. The distribution of the number of clues is notably different from the other samples.
How did you do the eliminations: single pass or two passes?

Back to top profile pm www

denis_berthier

Posted: Fri Jul 03, 2009 9:02 am Post subject: quote

Joined: 19 Jun 2007
Posts: 631
Location: Paris, France

Red Ed wrote:

> **denis_berthier wrote:**
>
> > Yes, but why 3322 patterns?
> > I guess they are not completely arbitrary.
> > Do they form some complete set of patterns wrt to some criterion?
>
> The historical records don't go back that far ... IIRC, it was just a collection of patterns that were good at defeating the generators of the day.

Thanks for this precision.
I've corrected my posts, in order to avoid confusions (mainly replacing "unavoidable sets" by "Red Ed's patterns").

Back to top profile pm www

m_b_metcalf

Posted: Fri Jul 03, 2009 10:03 am Post subject: quote

Joined: 15 May 2006
Posts: 2147
Location: Berlin

denis_berthier wrote:

> **m_b_metcalf wrote:**
>
> > a first result, if sustained, is tantalizing
> > ...
> > and includes an SE=9.1.
>
> Interesting. The distribution of the number of clues is notably different from the other samples.
> How did you do the eliminations: single pass or two passes?

This is a completely different program to my other one. This does a single pass to obtain a *locally* minimal puzzle. When I've checked everything, I may try out an *optimal* pass: my normal elimination removes redundant clues at random until the puzzle is minimal, but it is also possible to establish a list of all independently redundant clues and determine which combination of eliminations gives an *absolutely* minimal puzzle. I know from some earlier work that smaller puzzles can be obtained that way.

Regards,

Mike Metcalf

Back to top profile pm

denis_berthier

Posted: Fri Jul 03, 2009 10:21 am Post subject: quote

Joined: 19 Jun 2007
Posts: 631
Location: Paris, France

m_b_metcalf wrote:

> This does a single pass to obtain a *locally* minimal puzzle.

I'm mostly interested by this locally minimal version. The absolutely minimal one is likely to be biased (but interesting to see how much biased).
Regards

Back to top profile pm www

Red Ed

Posted: Fri Jul 03, 2009 10:36 am Post subject: quote

Joined: 06 Jun 2005
Posts: 522

denis_berthier wrote:

> ...
>
> **coloin wrote:**

> Perhaps the best way to get a random selection of puzzles is to just stick to 25 clue puzzles !
>
> ...
> I think sticking to 25 (or to any predefined number of) clues is one of the worst ways of getting a random sample of puzzles: it is an arbitrary and artificial constraint.

I beg to differ:

- It's already been shown that different generators can give rise to minimal puzzles with significantly different (distributions of) numbers of clues. I don't think anyone's really surprised by that, are they?
- Now it would be good to fix the number of clues and see if different generators can give significantly different (distributions of) complexity (SE, say) rating. It would be surprising and interesting if so, and worth trying to understand in detail.

Back to top profile pm

denis_berthier

Posted: Fri Jul 03, 2009 11:06 am Post subject: quote

Joined: 19 Jun 2007
Posts: 631
Location: Paris, France

> **Red Ed wrote:**
>
> Now it would be good to fix the number of clues and see if different generators can give significantly different (distributions of) complexity (SE, say) rating. It would be surprising and interesting if so, and worth trying to understand in detail.

As the number of clues is almost uncorrelated with the (SER or NRZT) complexity of puzzles, I don't think introducing this intermediate variable would be of much help for the global classification purposes of this thread. It can only make things more complex, especially as the distribution of the number of clues is not known a priori.

Back to top profile pm www

Red Ed

Posted: Fri Jul 03, 2009 11:14 am Post subject: quote

Joined: 06 Jun 2005
Posts: 522

You've analysed only means (average values), not whole distributions.

Do you really find uninteresting the possibility that the distribution of SE (say) ratings from suexg's 25-clue puzzles might differ from those from gsf's generator? I don't think such a result has been demonstrated yet, has it?

Back to top profile pm

denis_berthier

Posted: Fri Jul 03, 2009 11:41 am Post subject: quote

Joined: 19 Jun 2007
Posts: 631
Location: Paris, France

> **Red Ed wrote:**
>
> You've analysed only means (average values), not whole distributions.

Location: Paris, France

Blatantly false: I've given detailed classification results (SER and NRCZT distributions) for the full sudogen0_1M collection.
Moreover, on my web pages, for each puzzle in the sudogen0_1M collection you can find all the elements (#clues, SER, NRCZT, solution grids, ...) necessary to compute these distributions and anything you want for the puzzles with the desired number of clues.

> **Red Ed wrote:**
>
> Do you really find uninteresting the possibility that the distribution of SE (say) ratings from suexg's 25-clue puzzles might differ from those from gsf's generator?

It isn't the comparison I find uninteresting. It is the arbitrary restriction to 25. Why restrict this comparison to 25 (or any number of) clues and not make it for all the puzzles generated by these two generators?

> **Red Ed wrote:**
>
> I don't think such a result has been demonstrated yet, has it?

I'll be glad if you do it. As for me, I've never been able to find a version of gsf's program that I could run on my Mac or even a description of how it works.

Back to top profile pm www

Red Ed

Posted: Fri Jul 03, 2009 12:03 pm Post subject: quote

> **denis_berthier wrote:**
>
> > **Red Ed wrote:**
> >
> > You've analysed only means (average values), not whole distributions.
>
> Blatantly false: I've given detailed classification results ...

Joined: 06 Jun 2005
Posts: 522

I was referring to the comparison of different generators, but no matter.

The 25 clue restriction is just there to take a slice through the data, to avoid the conditional probabilities being washed away. 25's arbitrary, sure.

You (Denis) have kindly made available SE and NRCZT ratings for sudogen0_1M. Can you (or **anyone**) tell me where to find SE and/or NRCZT ratings, plus the puzzles themselves, for any **other** collection of minimal puzzles? Then I can do the test that I alluded to.

Back to top profile pm

Red Ed

Posted: Fri Jul 03, 2009 12:18 pm Post subject: quote

Further to that ^^^^ ...

Joined: 06 Jun 2005
Posts: 522

Can someone tell me how to:
(a) generate 1000 random minimal puzzles with gsf's code
(b) use gsf's code to rate a collection of puzzles

Back to top profile pm

denis_berthier

Posted: Fri Jul 03, 2009 9:27 pm Post subject: quote

Joined: 19 Jun 2007
Posts: 631
Location: Paris, France

> **Red Ed wrote:**
>
> Can you (or **anyone**) tell me where to find SE and/or NRCZT ratings, plus the puzzles themselves, for any **other** collection of minimal puzzles?

Not quite *any* collection? If you want to be able to reach some conclusion, you need at least randomness, no obvious bias (and therefore some clearly defined construction process) ....

Red Ed,
I can't be of any help for gsf's program. I'm not even sure it can output a random series of minimal puzzles.

For the SER, you can compute it with Sudoku Explainer 12.2.0. The (Unix) command line I use is:
java -cp SudokuExplainer.jar diuf.sudoku.test.serate --format=%r --input=puzzles.txt --output=ser.txt
where "puzzles.txt and "ser.txt" are the obvious file names.
After that, you just have to be patient.

For the NRCZT-ratings, currently only SudoRules can compute them. It takes more time than SER (it took me 5 weeks for 1,000,000 puzzles). I suggest you start with SER and, if the results look promising, I can do the NRCZT computations for you.

Back to top profile pm www

Red Ed

Posted: Fri Jul 03, 2009 11:06 pm Post subject: quote

Joined: 06 Jun 2005
Posts: 522

> **denis_berthier wrote:**
>
> > **Red Ed wrote:**
> >
> > Can you (or **anyone**) tell me where to find SE and/or NRCZT ratings, plus the puzzles themselves, for any **other** collection of minimal puzzles?
>
> Not quite *any* collection? If you want to be able to reach some conclusion, you need at least randomness, no obvious bias (and therefore some clearly defined construction process) ....

Indeed. I really just need to know how to generate and rate minimal puzzles. Thanks for help on rating; now I just want **gsf** to step in and tell me how to generate.

EDIT: scrub that, I've had a better idea. Sudogen0_1M is a big enough collection that we might learn something from comparing the complexity ratings of filtered subsets. For example: subset 1 = 23-clue puzzles; subset 2 = 27-clue puzzles (this has already been done of course). Another example: sort puzzles by how "unbalanced" their clue positions are; subset 1 = the top 10000, subset 2 = the

bottom 10000. If time permits later then I'll do some experiments.

Back to top profile pm

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 ... , 13, 14, 15 Next

Page 14 of 15

Jump to: Advanced solving techniques Go

You **cannot** post new topics in this forum
You **cannot** reply to topics in this forum
You **cannot** edit your posts in this forum
You **cannot** delete your posts in this forum
You **cannot** vote in polls in this forum