Sudoku Players' Forums

FAQ Search Memberlist Usergroups
Profile You have no new messages Log out [denis_berthier]

Rating rules / Puzzles. Ordering the rules

Goto page Previous 1, 2, 3 ... , 23, 24, 25 Next

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

View previous topic :: View next topic

Author	Message
StrmCkr Joined: 05 Sep 2006 Posts: 478 Location: Winterpeg	Posted: Thu Jul 09, 2009 7:20 am Post subject: quote **Quote:** deletion phase which part of the deletion phase the +2 - 1 phase or the after affect which has the affect of modify the logic sets the grid starts with. of removing super flours clues. (which merely reinforces the complexity of logic needed)
Back to top	profile pm
m_b_metcalf Joined: 15 May 2006 Posts: 2207 Location: Berlin	Posted: Thu Jul 09, 2009 7:23 am Post subject: quote **denis_berthier wrote:** 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 **Mike**, do you think your bottom-up generator could be adapted to do this? (10,000 couples of puzzles would be enough for such computations.)

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
```

I'll e-mail you the file.

Regards,

Mike Metcalf

Back to top

profile pm

denis_berthier

Joined: 19 Jun 2007
Posts: 678
Location: Paris, France

Posted: Thu Jul 09, 2009 7:29 am Post subject: quote edit

m_b_metcalf wrote:

> I added just three statements to my program, and got the following result in 500s

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.

Regards

Back to top

profile pm www

Allan Barker

Joined: 21 Feb 2008
Posts: 288
Location: Bangkok

Posted: Thu Jul 09, 2009 8:19 am Post subject: quote

> **denis_berthier wrote:**
>
> > **m_b_metcalf wrote:**
> >
> > I added just three statements to my program, and got the following result in 500s
>
> 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.
> Regards

I think this difference, which Mike pointed out seems to be ½ comes about purely from the direction you approach the boundary between MS and SS grids. Maybe people with more experience can help this one out.

When approaching from the top, the first candidate removed that produces a MS grid is replaced (put back in) and a local minimum is searched for. However, replacing this candidate effectively cuts the local sample space for this search in half by assuming this candidate must be in the minimum grid, which will be true 50% of the time (within a given distance in grid space). No such problem occurs coming up from the bottom because the first candidate added that produces a SS grid is assured of being in the local minimum. So, on average the top down approach will be higher (than necessary) by one candidate 50% of the time.

I have waived my hands, they are dry. ☹

Back to top

profile pm www

David P Bird

Joined: 17 Sep 2008
Posts: 60
Location: Middle England

Posted: Thu Jul 09, 2009 8:45 am Post subject: quote

I think what lies behind **StrmCkr's** thinking is something like this:

If we have supplied sufficient clues to solve a certain key part of a puzzle with difficulty, adding random clues to allow the rest of the puzzle to be completed would usually provide a much easier route to solve the key area. To prevent that, we would need to customise the balance of the givens to specifically cover the relatively unimportant fringe areas only.

On another subject, I have a gut feeling that the number and nature of the two-digit unconditional sets have a big role to play in the potential difficulty of a puzzle.

Here is one illustrative example in a grid with a single unconditional set for a digit pair:

Code:

```
|-------|-------|-------|
|     a | a     |       |
|       |   b   |     b |
| a     |       | a     |
|-------|-------|-------|
|       |     a |   a   |
|       |     a | a     |
| a a   |       |       |
|-------|-------|-------|
|   a a |       |       |
|       | a     |   a   |
|       |   b   |     b |
|-------|-------|-------|
```

The cells marked 'a' are all connected by an orthogonal loop with alternate legs along rows and columns. This leaves four further 'b' cells in a rectangle which from a "NOT a UR". If no given is provided for the 'b' cells, there is a chance that they won't be solvable until one of the 4 'a' cells in a common box has been resolved, which could restrict the solution path.

Back to top

profile pm

eleven

Joined: 10 Feb 2008
Posts: 331

Posted: Thu Jul 09, 2009 10:07 am   Post subject:

quote

One more puzzling result. I generated minimal puzzles bottom up without dropping any clues, i.e. i only took unique puzzles, if they were minmal from the adding clue phase.
Maybe Mike can repeat it to exclude coding errors.

Code:

```
20:      3
21:     46
22:    524
23:   2301
24:   3757
25:   2549
26:    710
27:    106
10000 puzzles, av. clue number 24.0804
```

To my surprise the average clue number only is slightly higher and still clearly under the value from top down generation.

Back to top

profile pm

m_b_metcalf

Joined: 15 May 2006
Posts: 2207
Location: Berlin

Posted: Thu Jul 09, 2009 10:19 am   Post subject:

quote

eleven wrote:

> Maybe Mike can repeat it to exclude coding errors..

After the adding-clue phase I typically have 30-40 clues. A quick run gave zero minimal puzzles.

Regards,

Mike Metcalf

Back to top profile pm

eleven

Joined: 10 Feb 2008
Posts: 331

Posted: Thu Jul 09, 2009 10:30 am Post subject: quote

> **m_b_metcalf wrote:**
>
> > **eleven wrote:**
> >
> > Maybe Mike can repeat it to exclude coding errors..
>
> After the adding-clue phase I typically have 30-40 clues. A quick run gave zero minimal puzzles.

It took some hours to get the 10000 puzzles.

Back to top profile pm

coloin

Joined: 06 May 2005
Posts: 1017
Location: Oxford

Posted: Thu Jul 09, 2009 11:13 am Post subject: quote

I think the decreased clues in the "bottom up" method can be explained.

The number of puzzles within a {-1+1} or {-2+2} is greater in easier puzzles than in harder puzzles. Therefore when you are picking clues to complete the puzzle there will be more easy puzzles.

Why might this be...?

An easy [singles] puzzle if you remove a clue - still partially completes with singles - and the n-1 subpuzzle will tend to have less "other grid solutions". Less "other grid solutions" implies that there will be more {+1} ways to complete the puzzle. All the puzzles generated from the {n-1} subpuzzle will have the singles which were in the subpuzzle. So all will tend to be easy.

I have long thought about the analysis of the {n-1} subpuzzles in "hard" puzzles.Perhaps the lack of infered clues in ALL the n-1 subpuzzles tends to imply a hard puzzle.

C

Back to top profile pm

Allan Barker

Joined: 21 Feb 2008
Posts: 288
Location: Bangkok

Posted: Thu Jul 09, 2009 1:36 pm Post subject: quote

> **Allan Barker wrote:**
>
> > **denis_berthier wrote:**
> >
> > > **m_b_metcalf wrote:**
> > >
> > > I added just three statements to my program, and got the following result in 500s
> >
> > 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.
> > Regards
>
> I think this difference, which Mike pointed out seems to be ½ comes about purely from the direction you approach the boundary between MS and SS grids.

I have some results for the following

A): How to make a top down generator produce (Edit) average clue sizes of a bottom up generator.
B): Why the top down and bottom up generators differ by ½.

These experiments are all done with a top down generator.

Precursor = P: All experiments include: (random 81 clue grid) + top down until removal of A produces a multi-solution grid.

Experiment 1 P + Re-place A, minimize puzzle.
Experiment 2 P + don't re-place A, add one random clue B to make valid puzzle, minimize puzzle.
Experiment 3 Perform both exp 1 and exp 2 on the same puzzle, Choose smaller of the two minimal puzzles.

Experiment 1 is as per normal for a top down approach.
Experiment 2 produces a second version of the same puzzle with same number of clues that differs only by A versus B.
Experiment 3 Combining exp 1 and exp 2 is the same process with double the effective search space for a minimal puzzle

Explanation: Re-placing A in the top down approach effectively cuts the search space for a minimal puzzle in half (one clue of N) relative to a bottom up approach, (see previous post above) Doubling the search space in experiment 3 should then produce results (eg, average number of clues) similar to a bottom up generator.

Edit: As noted by Eleven below, although Exp 3 produces the same mean clue size as the bottom up, the ditribution is tighter, i.e., the there are 5% more 24's (41.6 to 36.4)

Results: Adding clue B instead of clue A should produce no difference from replacing A.
Combining A and B doubles the effective search area and should thus produce similar mean clue values to a bottom up generator.

Code:

```
C 19        0        Experiment 1
C 20        1        Re-place A, minimize puzzle
```

```
C 21      43
C 22     694
C 23    3379
C 24    6886
C 25    5983
C 26    2440
C 27     497
C 28      70
C 29       7
C 30       0
N=20000 avg = 24.3885

C 19       0      Experiment 2
C 20       1      Don't re-place A
C 21      47             Find 1 random clue B to make
valid grid. minimize puzzle
C 22     740
C 23    3472
C 24    6837
C 25    5884
C 26    2463
C 27     505
C 28      46
C 29       5
C 30       0
N=20000 avg = 24.3733

C 19       0          Combine Exp 1 + Exp 2,
C 20       0     Choose smallest minimal puzzle
C 21      72
C 22    1366
C 23    5834
C 24    8318
C 25    3821
C 26     556
C 27      33
C 28       0
C 29       0
C 30       0
N=20000 avg = 23.8125
```

Last edited by Allan Barker on Fri Jul 10, 2009 1:36 am; edited 1 time in total

Back to top profile pm www

StrmCkr

Joined: 05 Sep 2006
Posts: 478
Location: Winterpeg

Posted: Thu Jul 09, 2009 2:47 pm Post subject: quote

Quote:

> Results: Adding clue B instead of clue A should produce no difference from replacing A.
> Combining A and B doubles the effective search area and should thus produce the same results as the bottom up generator.

adding clue B can change the # of minimized clue...

doing this changes the preset logic bases and should effectively mimic
a bottom up generator as minimality is no longer depend on the preset logic bases from the starting grid.

david
thanks you for the rewrite

yes that was my point exactly in bottom up generators it has the tendency to fill in a key area randomly and prevent complex logic sets.

in top down the logic created by the starting grid aid by staying constant where by checking the solution count u can determin if you can still remove a clue or not.

Back to top profile pm

m_b_metcalf

Joined: 15 May 2006
Posts: 2207
Location: Berlin

Posted: Thu Jul 09, 2009 3:03 pm Post subject: quote

eleven wrote:

> **m_b_metcalf wrote:**
>
> > **eleven wrote:**
> >
> > > Maybe Mike can repeat it to exclude coding errors..
> >
> > After the adding-clue phase I typically have 30-40 clues. A quick run gave zero minimal puzzles.
>
> It took some hours to get the 10000 puzzles.

Indeed. So I made a short run:

Code:

```
Number:          157
Average:    24.10191

      21       0
      22       7
      23      29
      24      57
      25      46
      26      15
      27       1
      28       1
      29       0
```

Looks comparable?

Mike Metcalf

Back to top profile pm

coloin

Joined: 06 May 2005
Posts: 1017

Posted: Thu Jul 09, 2009 3:11 pm Post subject: quote

Yes, in the bottom up there is an advantage in being able to pick a [unique] clue [from all the grid solutions of the subpuzzle] which gives a valid puzzle.

Location: Oxford

They will be tend to be easier puzzles if there are infered clues [singles] already in the subpuzzle - which they tend to be if the subpuzzle is able to be completed to puzzle with one more clue..

This disadvantages the top down method where the solution grid is constant.

C

Back to top profile pm

eleven

Joined: 10 Feb 2008
Posts: 331

Posted: Thu Jul 09, 2009 7:02 pm Post subject: quote

Allan Barker wrote:

> However, replacing this candidate effectively cuts the local sample space for this search in half by assuming this candidate must be in the minimum grid, which will be true 50% of the time (within a given distance in grid space).

I dont understand this. The top down generation is nothing than looking for a local minimum in a whole grid, done by eliminating clues in random order. Some of them would lead to multi solution puzzles, so they are not eliminated.

Now you say, when you arrive at the first number, which cannot be eliminated, you replace it by another one, thus building local minima 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 ?

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).

Anyway, it was another surprise for me, that you could lower the mean value that much with this simple change.

Back to top profile pm

Red Ed

Joined: 06 Jun 2005
Posts: 549

Posted: Thu Jul 09, 2009 9:11 pm Post subject: quote

This thread seems to be getting a bit mixed up.

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. (Not trying to be Thought Police here: just genuinely confused.)

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 ...

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

Goto page Previous 1, 2, 3 ... , 23, 24, 25 Next

Page 24 of 25

Stop watching this topic

Jump to: Advanced solving techniques Go