Sudoku Players' Forums

THE REAL DISTRIBUTION OF MINIMAL PUZZLES

Goto page Previous 1, 2, 3 ... , 40, 41, 42 Next

newtopic postreply **Sudoku Players' Forums Forum Index -> General/puzzle**

View previous topic :: View next topic

Author	Message
JPF Joined: 07 Dec 2005 Posts: 2965 Location: Paris, France	Posted: Mon Oct 19, 2009 11:23 am Post subject: quote Just a naive observation. Sudoku Explainer and SER make the assumption that the puzzles have only one solution, and therefore use all the UR, BUG techniques... Let SER* be the SER without these techniques. I guess that SER*>= SER What would be the correlation between SER* vs. NRCZT or vs. n (n being the number of clues)? PS: Here are the SER ratings : 1.0: Last value in block, row or column 1.2: Hidden Single in block 1.5: Hidden Single in row or column 1.7: Direct Pointing 1.9: Direct Claiming 2.0: Direct Hidden Pair 2.3: Naked Single 2.5: Direct Hidden Triplet 2.6: Pointing 2.8: Claiming 3.0, 3.2, 3.4: Naked Pair, X-Wing, Hidden Pair 3.6, 3.8, 4.0: Naked Triplet, Swordfish, Hidden Triplet 4.2, 4.4: XY-Wing, XYZ-Wing **4.5 - 5.0: Unique rectangles and loops** 5.0, 5.2, 5.4: Naked Quad, Jellyfish, Hidden Quad **5.6 - 6.0: Bivalue Universal Graves** 6.2: Aligned Pair Exclusion 6.5 - 7.5: Bidirectioal X-Cycles and Y-Cycles 6.6 - 7.6: Forcing X-Chains 7.0 - 8.0: Forcing Chains, Bidirectional Cycles 7.5 - 8.5: Nishio 8.0 - 9.0: Cell/Region Forcing Chains 8.5 - 9.5: Dynamic Forcing Chains 9.0 - 10.0: Dynamic Forcing Chains (+) > 9.5: Nested Forcing Chains

JPF

Back to top profile pm

denis_berthier

Posted: Mon Oct 19, 2009 11:46 am Post subject: quote edit

Joined: 19 Jun 2007
Posts: 955
Location: Paris, France

JPF wrote:

> Just a naive observation.
> Sudoku Explainer and SER make the assumption that the puzzles have only one solution, and therefore use all the UR, BUG techniques...
> Let SER* be the SER without these techniques.
> I guess that SER*>= SER

SER can be criticised in many respects.
But it is the case that it is well correlated (~ 0.9) with the better NRCZT rating (purely logical, invariant under super-symmetries, but longer to compute) .
I use it only as a first approximation of NRCZT.
For your specific question about uniqueness techniques, I don't think it would change much (wrt its correlation with NRCZT or to its own distribution) to eliminate them.
If I knew how to do this, I could try on a small sample.

Last edited by denis_berthier on Mon Oct 19, 2009 4:59 pm; edited 1 time in total

Back to top profile pm www

Allan Barker

Posted: Mon Oct 19, 2009 12:18 pm Post subject: quote

Joined: 21 Feb 2008
Posts: 357
Location: Bangkok

denis_berthier wrote:

> I have a problem with your 34s file. I can't read the file after the 385th puzzle.

Fixed. There was a stray end of file character at that point.

denis_berthier wrote:

> .... if you could generate 1000 puzzles with 25, 26 or 27 clues, we could compare.

OK, that's easy enough to do.

Back to top profile pm www

Pat

Posted: Mon Oct 19, 2009 2:43 pm Post subject: quote

Joined: 18 Jul 2005
Posts: 1575

denis_berthier wrote:

> If I knew how to do this, I could try on a small sample.

you'd need to make a small revision in the source-code
(to disable those specific techniques)

and compile this special-purpose version---

Back to top profile pm

denis_berthier

Posted: Tue Oct 20, 2009 8:01 am Post subject: quote edit

Joined: 19 Jun 2007
Posts: 955
Location: Paris, France

HOW THE SER REALLY DEPENDS ON THE NUMBER OF CLUES

In a previous post:
http://www.sudoku.com/boards/viewtopic.php?t=14615&postdays=0&postorder=asc&start=554,
I gave an estimate of the real number-of-clues distribution of mininal puzzles. It was based on a sample of 1,380,962 minimal puzzles, generated with the controlled-bias generator, taking as input 65 full scans of gsf's collection of all the (equivalence classes of) complete grids.

I also gave a first, rough estimated value for the real SER (mean and standard deviation), in which the mean value and standard deviation for each fixed number of clues were taken from those computed for the Sudogen0_1M collection.

Now, I have finished computing the SER for all the above 1,380,962 puzzles and I thus have much better estimates for each number of clues and for the global values.
Remember that, for any fixed number of clues, the controlled-bias generator, when it uses an integer number of full scans of gsf's collection, is completely unbiased. As a result, each row of the the table below gives both the controlled-bias and the real values for the n-clue SER. Only the global mean values and standard deviations have to be computed differently (without or with the correction coefficients).

Code:

```
#clues  #instances       mean(SER)       standard-
deviation(SER)
19      0
20      0
21      41              3.56 (*)        2.01 (*)
22      1,526           3.15            2.16
23      25,884          3.35            2.24
24      163,694         3.61            2.36
25      422,451         3.96            2.47
26      467,047         4.40            2.54
27      234,963         4.93            2.53
28      57,615          5.47            2.44
29      7,243           6.07            2.19
30      481             6.76            1.71
31      16              5.79 (*)        2.34 (*)
32      1               7.3 (*)         (*)
all    1,380,962

(*) values based on small samples are not meaningful.
```

Which gives:

Code:

```
controlled-bias mean(SER) = 4.29    controlled-bias
standard-deviation(SER) = 2.48
real mean(SER) = 4.73               real standard-
deviation(SER) = 2.49
(These figures are the same when suexg-cb is used as
the source of complete grids)

correlation coefficient #clues vs SER = 0.19
```

What's most noticeable, when we compare with the results obtained for the various generators (bottom-up, top-down, full bottom-up), is **a stronger upward complexity trend wrt the number of clues**.
In the [22, 29] range where they can be compared (enough instances in the samples), the mean starts from smaller SER (3.15 for 22 clues vs 3.27 for top-down) and reaches higher SER (6.07 for 29 clues vs 5.43 for top-down). But the trend in the above results can be seen beyond that interval. It may also be opportune to recall that we probably know a large proportion of the 17-clue minimals, for which the mean SER is 2.55.

The correlation coefficient #clues vs SER (= 0.19) is a little higher than for the top-down generator (0.12) but it remains too small to allow any predictions of complexity given the number of clues.

If needed, this confirms, once more and much more strongly than before, what I've said many times:
- the statistics for the (SER or NRCZT) complexity of puzzles depends strongly on the way these puzzles are built;
- in particular, building collections of puzzles that are unbiased wrt to the number of clues is no guarantee that they are unbiased wrt (SER or NRCZT) complexity.

Last edited by denis_berthier on Wed Oct 21, 2009 5:03 am; edited 1 time in total

Back to top profile pm www

Red Ed

Posted: Tue Oct 20, 2009 4:00 pm Post subject: quote

Joined: 06 Jun 2005
Posts: 826

To help you draw out the trend further, I'll get around to computing the mean SER for my 700+ 32s at some point (EDIT: done - it's 6.84 over 711 puzzles). I'm surprised that you've not already included **eleven**'s result, i.e. mean 6.88 over 369 clustered puzzles.

It would be good to see the same results for NRCZT. With that in mind, can I draw your attention to my recent "algorithm" thread?

Back to top profile pm

denis_berthier

Posted: Wed Oct 21, 2009 5:17 am Post subject: quote edit

Joined: 19 Jun 2007
Posts: 955
Location: Paris, France

COMPLEMENTS TO MY PREVIOUS POST

When suexg-cb is used as the source of complete grids (instead of gsf's collection, as in my previous post), I had forgotten to publish the following results.

The SER mean and standard deviations (based on 350,000 puzzles) are the same as for the gsf source.

The correlation coefficients are:
#clues vs SER = 0.20
#clues vs NRCZT = 0.19
SER vs NRCZT = 0.90

For the NRCZT (based on 250,000 puzzles):

Code:

```
controlled-bias average = 2.22                real
(estimated) average = 2.45
controlled-bias standard-deviation = 1.35     real
(estimated) standard-deviation = 1.39
```

As a result, there's no reason to suppose that anything will be changed for these NRCZT values when gsf is used as the source of complete grids and I don't plan to spend time to make additional NRCZT computations for this case.

Back to top

profile pm www

Red Ed

Posted: Wed Oct 21, 2009 6:15 am Post subject:

quote

Joined: 06 Jun 2005
Posts: 826

It's a pity that you won't do the NRCZT calculation for gsf's collection, as it would be interesting to observe whether or not the "more clues => harder" trend is reflected there as well. Although I regard NRCZT chains' restriction that Rn is weak-linked to Ln+1 as a somewhat arbitrary, it's not nearly as arbitrary as SE's ordering of techniques, and so the trend in NRCZT chain ratings would be more compelling IMO. An NRCZT braid rating would be even better.

Back to top

profile pm

denis_berthier

Posted: Wed Oct 21, 2009 6:43 am Post subject:

quote edit

Joined: 19 Jun 2007
Posts: 955
Location: Paris, France

Red Ed wrote:

> It's a pity that you won't do the NRCZT calculation for gsf's collection, as it would be interesting to observe whether or not the "more clues => harder" trend is reflected there as well.

You should re-read my last post. For the SER, there's no difference between the results obtained with the 2 sources of complete grids: gsf's source or suexg's internal source.

In gsf's source, the trend is present for the SER. Due to the strong correlation SER vs NRCZT, the trend is there for the NRCZT also. So there's no need to waste time on something that wouldn't bring any new insight.

> **Red Ed wrote:**
>
> Although I regard NRCZT chains' restriction that Rn is weak-linked to Ln+1 as a somewhat arbitrary.

This "arbitrary" restriction on nrczt-chains/whips is what makes them chains instead of nets. When you build an nrczt-chain/whip, you follow a single line of thought, contrary to what you have to do with a net.
nrczt-braids are nets (not the worst kind one can imagine, but nets anyway). *I* don't like nets. The main interest of nrczt-braids is purely theoretical: it is my braids vs T&E theorem.
But, if *you* like nets, I have no objection.

Nevertheless, this "arbitrary" restriction makes my results much stronger than if I had stated them for braids: all the puzzles generated by random generators can be solved by nrczt-whips (and I have effectively solved several millions).

Back to top profile pm www

Red Ed

Posted: Wed Oct 21, 2009 7:25 am Post subject: quote

Joined: 06 Jun 2005
Posts: 826

IIRC, you've said in the past something along the lines of: SER is of interest only as a relatively cheap-to-calculate approximation of the NRCZT rating. Paul's given you the means to perform the NRCZT rating calculation directly, very quickly, thereby doing away with the need for approximation using SE. Surely when illustrating the thesis that "more clues => harder" it would be better to do so using the rating of greater interest rather than an approximation? I don't see why you don't just fire off Paul's code and harvest the results.

btw, as you can tell, I am catching up on the NRCZT story. Where's the braids vs. T&E theorem?

Back to top profile pm

PIsaacson

Posted: Wed Oct 21, 2009 8:16 am Post subject: quote

Joined: 02 Jul 2008
Posts: 241
Location: Campbell, CA

Red Ed,

Without some major modifications to my code, it can't be counted on to find the smallest nrczt chain required to solve any particular puzzle. I don't execute an incremental depth limit on the BFS, so in it's current state, it may or may not find an equivalent shortest solving chain. I would need to coordinate with Denis (not that I wouldn't give my eye-teeth to do so, hint hint!!!) on many design details in order to replicate what his SudoRules accomplishes.

I'm already ripping apart my code and putting together a callable library, but there are lots of issues to resolve before it can be used to score a puzzle.

Cheers,
Paul

Back to top profile pm

denis_berthier

Posted: Wed Oct 21, 2009 8:28 am Post subject: quote edit

Joined: 19 Jun 2007
Posts: 955
Location: Paris, France

Red Ed wrote:

> IIRC, you've said in the past something along the lines of: SER is of interest only as a relatively cheap-to-calculate approximation of the NRCZT rating. Paul's given you the means to perform the NRCZT rating calculation directly, very quickly, thereby doing away with the need for approximation using SE.

See Paul's answer. Sure I'd like to have a fast NRCZT rater.
But, in this specific case, the insensitivity of the controlled-bias generator to the source of complete grids gives the answer.
I still have to publish the suexg-cb NRCZT distribution (no long computations to do).

Red Ed wrote:

> Where's the braids vs. T&E theorem?

"Abominable T&E vs lovely braids" thread.
Or my web pages.

Back to top profile pm www

denis_berthier

Posted: Wed Oct 21, 2009 8:44 am Post subject: quote edit

Joined: 19 Jun 2007
Posts: 955
Location: Paris, France

PIsaacson wrote:

> I would need to coordinate with Denis (not that I wouldn't give my eye-teeth to do so, hint hint!!!) on many design details in order to replicate what his SudoRules accomplishes.

Replicating the details is not necessary. Indeed, you have no chance of finding exactly the same paths, as there is a random choice.
What's important is finding the same ratings. Sudogen0_1M (see my web pages) provides 1,000,000 puzzles against which you can test your software.
I can only encourage you to issue a version of your software for fast NRCZT rating. That'd be very useful to all of us, even if it doesn't print correctly the solution path.
I've already proven that, as far as only the rating is concerned, NRCZT can be approximated with B-NRCZT (based on nrczt-braids instead of whips), so that you can choose whatever's most convenient for you.

As for "design details", we've already evoked the question. SudoRules has no code in the traditional sense: the nrczt rules are written in CLIPS syntax, i.e. almost as pure logic formulæ. Any optimisations I've done are very specific to inference engines and have no procedural counterpart.
Unfortunately, for C, Pascal, Java or whatever procedural language you chose for your implementation, I'm almost totally incompetent.

But if you have any questions that you can express in natural language, I'll answer (preferably in the "nrczt-chains" or the "rating" thread, as this has nothing to do with the distribution problem).

Back to top profile pm www

Allan Barker

Posted: Wed Oct 21, 2009 1:28 pm Post subject: quote

Joined: 21 Feb 2008
Posts: 357
Location: Bangkok

Denis Berthier wrote:

> I've also thought of a way of testing your algorithm for bias. Now that we know precisely the mean SER for every n in [22, 30] (I'll soon publish my results, based on ~ 1,300,000 puzzles): if you could generate 1000 puzzles with 25, 26 or 27 clues, we could compare.

Denis,

I have prepared a new series of random minimal puzzles, which should make for an interesting comparison to the rest of the data. I have also extended the upper range with 500 new 35-clue minimals and (with no concern for global warming) 20 **36-clue** minimals. All puzzles were made the same way and are random/unrelated. I would assume that the bias is roughly that of other puzzles generated by similar means.

I extended the range downwards in case it helps with interpreting the overlap region.

Code:

```
  size   count
---------------
Size 22  1000
Size 23  1000
Size 24  1000
Size 25  1000
Size 26  1000
Size 27  1000
Size 28  1000
Size 29  1000
Size 30  1000
Size 31  1190
Size 32  2610
Size 33  1000
Size 34  1000
Size 35   500
Size 36    20
```

PS. No end of file characters this time.

Back to top profile pm www

denis_berthier

Posted: Thu Oct 22, 2009 5:47 am Post subject: quote edit

Allan,

Joined: 19 Jun 2007

Posts: 955
Location: Paris, France

Thanks for these data.
A quick SER computation gives:

Code:

```
#c  mean  sd
22  3.20  2.13
23  3.18  2.16
24  3.58  2.36
25  3.98  2.48
26  4.31  2.55
27  4.79  2.54
28  5.34  2.51
29  5.97  2.28
30  6.32  2.14
31  6.67  1.86
32  6.89  1.71
33  7.14  1.49
34  7.25  1.44
35  7.54  1.18
36  7.43* 0.94

* too few data to be meaningful
```

We can see a very strong upward trend.

In the range where comparisons are possible, it is closer to the real values than any upward or top-down generator. Though systematically below, it is consistent with sampling errors.

What happens if you don't a preset the number of clues?

Back to top profile pm www

Display posts from previous: All Posts Oldest First Go

newtopic postreply **Sudoku Players' Forums Forum Index -> General/puzzle**

All times are GMT
Goto page Previous 1, 2, 3 ... , 40, 41, 42 Next

Page 41 of 42

Stop watching this topic

Jump to: General/puzzle Go