Sudoku Players' Forums

THE REAL DISTRIBUTION OF MINIMAL PUZZLES

Goto page Previous 1, 2, 3, 4, 5 Next

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

View previous topic :: View next topic

Author | **Message**

denis_berthier

Posted: Sun Jul 12, 2009 3:04 pm Post subject: quote edit

Joined: 19 Jun 2007
Posts: 712
Location: Paris, France

Red Ed wrote:

> **denis_berthier wrote:**
>
> > What happens below B (i.e. after a minimal puzzle is reached) is irrelevant. All that we need are the probabilities on B and above.
>
> It is highly relevant. Your **P24** formula, for example, assumes that all choose(81,24) subgrids at level 24 are "reachable" ... but they're *not* - e.g. because some have multiple solutions. But you can't say that level 24 is "below B" and thus "irrelevant" because there are minimal puzzles at that level that whose probabilities we need to know.

The P24 formula concerns only the forest of indexed puzzles, a static structure, whose existence is independent of any process that can occur on it.
Obviously, 24 is below B along some paths and above B along some other paths. Nothing's fixed in advance.
And there can't be minimal puzzles below B.
Obviously, you don't really read my answers.

Red Ed wrote:

> I still maintain that your **Pn** formula is out by a factor of s(n)/eps(n).

You're turning into circles by trying to introduce lots of new functions. All that's necessary is already in my previous posts. Needless to mess it up with irrelevant functions.

Back to top profile pm www

denis_berthier

Posted: Sun Jul 12, 2009 3:10 pm Post subject: quote edit

Joined: 19 Jun 2007
Posts: 712
Location: Paris, France

Red Ed wrote:

> **coloin wrote:**
>
> > You would think that statistically the chances of a [minimal] 24-puzzle being produced would be the same as any other [minimal] 24-puzzle - but I believe this is not the case.
>
> "Not the case" is correct. I didn't want to raise that myself in discussion with Denis because he talks about variables X(n) obtained by averaging over all n-clue puzzles in a sample; his focus appears not to be on the variance among individual puzzles.

Wrong.
The chances of any two minimal 24-puzzles are the same.
The variance of individual puzzles is irrelevant at that point.

Back to top profile pm www

Red Ed

Posted: Sun Jul 12, 2009 4:10 pm Post subject: quote

You just can't believe you're wrong, can you?

Joined: 06 Jun 2005
Posts: 582

I hereby give up trying to help you come to terms with your mathematical failings.

Back to top profile pm

denis_berthier

Posted: Sun Jul 12, 2009 4:20 pm Post subject: quote edit

Joined: 19 Jun 2007
Posts: 712
Location: Paris, France

> **Red Ed wrote:**
>
> You just can't believe you're wrong, can you?
> I hereby give up trying to help you come to terms with your mathematical failings.

If this can make you feel better...
But do you think this can make you right?
All the probability thing on the set of indexed puzzles is secondary school statistics, which was obviously not your major.

I think I now understand why it is so necessary for you to throw some doubt on my approach. I had a look at the other thread, where you claim that the mean number of clues is above 26. Unfortunately, your claim is totally unsubstantiated: you're referring only vaguely to computations you've made, but you have no theory at all and you give no indication on these computations.

But thanks anyway. Your nitpicking has helped me improve the redaction. Final version here:
http://www.carva.org/denis.berthier/HLS/Classification/index.html

Back to top profile pm www

Red Ed

Posted: Sun Jul 12, 2009 5:26 pm Post subject: quote

Joined: 06 Jun 2005
Posts: 582

I can't wait until some other mathematically-qualified people (apart myself: not you obviously) look at that web page. It's a morass of bloated verbiage, written more in the style of management theory than mathematics, based on foundations of sludge. Large chunks of it are so waffly and imprecise as to be, like that wonderful jibe at string theory, "not even wrong". So bad luck Denis; you've killed what could have been a constructive analysis by being impossible to work with.

Back to top profile pm

denis_berthier

Posted: Sun Jul 12, 2009 5:51 pm Post subject: quote edit

Joined: 19 Jun 2007
Posts: 712
Location: Paris, France

> **Red Ed wrote:**
>
> I can't wait until some other mathematically-qualified people (apart myself: not you obviously) look at that web page. It's a morass of bloated verbiage, written more in the style of management theory than mathematics, based on foundations of sludge. Large chunks of it are so waffly and imprecise as to be, like that wonderful jibe at string theory, "not even wrong". So bad luck Denis; you've killed what could have been a constructive analysis by being impossible to work with.

Because you think you are mathematically qualified, when you don't even understand secondary school statistics and your production in the other web page doesn't rely on any theory at all?

Your analysis has never been constructive. Since the beginning, your only goal was to try to discredit what I had written. Any means for this were good.
But, in spite of all your efforts to transform my theory into sludge by substituting your own notations and definitions to mine, you've been unable to find anything wrong in it. You keep repeating the same claims, without any rational argument.

Missing any rational argument, you now turn to insults.

Good luck, Red Ed.

Back to top profile pm www

Red Ed

Posted: Sun Jul 12, 2009 6:16 pm Post subject: quote

I turn to insults, regrettably, through sheer frustration at your incompetence.

Joined: 06 Jun 2005
Posts: 582

Re "no theory at all" - put your comments on the other thread where I can expose them for their vacuity.

Back to top profile pm

denis_berthier

Posted: Sun Jul 12, 2009 6:26 pm Post subject: quote edit

> **Red Ed wrote:**
> I turn to insults, regrettably, through sheer frustration at your incompetence.

Joined: 19 Jun 2007
Posts: 712
Location: Paris, France

More insults.
Sure you're showing here your main real competence.

> **Red Ed wrote:**
> Re "no theory at all" - put your comments on the other thread where I can expose them for their vacuity.

You'd better spend your time trying to give the unjustified assertions you make there the beginning of a shadow of explanation.
I won't waste my time on such vacuity.

Last edited by denis_berthier on Sun Jul 12, 2009 6:30 pm; edited 1 time in total

Back to top profile pm www

Red Ed

Posted: Sun Jul 12, 2009 6:26 pm Post subject: quote

Let's play "who can have the last word" 

Joined: 06 Jun 2005
Posts: 582

Back to top profile pm

denis_berthier

Posted: Sun Jul 12, 2009 6:31 pm Post subject: quote edit

> **Red Ed wrote:**
> Let's play "who can have the last word" 

Joined: 19 Jun 2007
Posts: 712
Location: Paris, France

You win

Back to top profile pm www

coloin

Posted: Sun Jul 12, 2009 9:14 pm Post subject: quote

Well, I have to say, before I go out to the pub to celebrate England's draw [], that, historically **Red Ed** isnt often wrong...[I think once he was].....and he has been a very valuable contributor to our understanding of grids and puzzles.

Joined: 06 May 2005
Posts: 1031
Location: Devon UK

My inital tests do indeed show that despite what you wold initially think [as I did] not all [same size minimal] puzzles are generated as frequently as others - possibly for the reasons I mentioned.

I generated all the different puzzles from a 40 clue subgrid [non-minimal puzzle].[using **gsf**'s program. There were ~1.5M puzzles, There were 3 with 20 clues.

In attempting to generate from the subgrid with suex* duplicate puzzles especially with low clues are generated.

Initial results show that indeed 1 of the puzzles was generated x2 more than the other 2.

I will perhaps do a more detailed test on this - but for now perhaps I will try "Ardbeg" Whisky.

Cheers

C

Back to top profile pm

coloin

Joined: 06 May 2005
Posts: 1031
Location: Devon UK

Posted: Mon Jul 13, 2009 12:26 am Post subject: quote

Denis I estimated mean puzzle size here

Essentially I processed all the different minimal puzzles from a 40 clue subgrid - I proposed a "fudge factor" based on the error - essentially the same as you have just done - although perhaps my method was fairly crude.

I worked out this fudge factor with 24 clues as the reference.here

I estimated what a 24.38 distibution would become.....

Code:

```
clues         ff           suexg       tot clues       new distribution

22           0.125          33966          93406            4245
23           0.38          170727        1492153           64876
24           1             342620        8222880          342620

25           2.2           298349       16409195          656367
26           4.5           122691       14354847          552109
27           7.6            25237        5178632          191801

28           8.5             2733         650454           23230



                           24.38       46401568         1835251
Average size 25.28
```

This is approaching your estimate of 25.39, although I was aware that I was underestimating. I could easily have compared the real distribution to a very sparse suex* - say up untill the first duplicate emergence.

You chose 20 clues as your reference, since this is towards the inaccurate end of the puzzle gen process, is it not more sensible to chose 24.....

I must say though that 26.5 seem a very high mean ..... 

C

Back to top profile pm

denis_berthier

Joined: 19 Jun 2007
Posts: 712
Location: Paris, France

Posted: Mon Jul 13, 2009 6:24 am Post subject: quote edit

coloin wrote:

> hstorically Red Ed isnt often wrong...[I think once he was].....and he has been a very valuable contributor to our understanding of grids and puzzles.

I have no doubt on this.
But even the best ones can make errors.

coloin wrote:

> Average size 25.28
> This is approaching your estimate of 25.39

Yes, very close. Your informal analysis was essentially correct.

> **coloin wrote:**
>
> You chose 20 clues as your reference, since this is towards the inaccurate end of the puzzle gen process, is it not more sensible to chose 24.....

The reference chosen is irrelevant. If I had chosen 24 instead of 20, the cf-corrections would have be different (all divided by the same number), but their ratios wouldn't have changed, nor would the unbiased means, in particular the 25.39 estimate.

I'll now prove in several easy steps (secondary school level) that all the minimal puzzles with the same number of clues have the same probability of being reached by a top-down generator.

Let N be the (huge) number of complete grids.
Let mmax be the maximum number of clues a minimal puzzle can have. We don't know the value of mmax, but we only need to know its existence.

step 1) all the puzzles (minimal or not) with mmax clues have the same probability of being reached by a top-down generator.
proof: as, at every step, the generator deletes one of the remaining clues with equal probabilities, for any m >= mmax, each puzzle with m clues has probability 1/N * 1/81! * (81 - m)! * m! of being reached.

Before I go further, do you agree with this ?

Back to top profile pm www

eleven

Posted: Mon Jul 13, 2009 8:01 am Post subject: quote

Joined: 10 Feb 2008
Posts: 346

> **denis_berthier wrote:**
>
> secondary school level

Thats for me

> **Quote:**
>
> step 1) all the puzzles (minimal or not) with mmax clues have the same probability of being reached by a top-down generator.
> proof: as, at every step, the generator deletes one of the remaining clues with equal probabilities, for any m >= mmax, each puzzle with m clues has probability 1/N * 1/81! * (81 - m)! * m! of being reached.

Please continue.

Back to top profile pm

coloin

Posted: Mon Jul 13, 2009 8:28 am Post subject: quote

Ok.....the chances of getting to the same number of clues would appear to be the same.

Joined: 06 May 2005
Posts: 1031
Location: Devon UK

Before you do go on consider this 40-clue subgrid [all clues are superfluous]

Code:

```
1....6..945.78....7.9.2354.2.4.9.37.6..57.914.75..1.....12.4.57..2.......6.935421

and its 11 20-puzzles within, with their freq. #, out of 353 total puzzles.

........94...8....7....35..2......7.....7.914..5........12.4.............6.9.5..1# 69
.....6...4..78......9...54.2.....3........91..75..1.....12....7..........6..35...# 17
.....6...4..78......9...54.2...9.3........9...75..1.....12....7..........6..35...#
```

```
30
1.......9.5..8........23.....4...3.....5..91..7...........4.57..2.......6.9.5..1#
26
1.......94...8....7....35..2.4....7.....7.91...5.........2.4.............6.9.5..1#
46
1.......94...8....7...235.........7.....7.914..5...........4.....2.......6.9.5..1#
48
1....6....5.........9.23.4...4.9.3..6..57.....7...1...........7..2.........9..4.1#
31
1....6....5.7.......9.2..4...4.9.3..6..5......7...1..........57..2..........3.42.#
25
1....6...4...8......9...5....4.9.3.....5..9...7...1......2.4..7..2.......6...5..1#
17
1....6...4...8......9.2.5......9.3..6..5..9...75..1...........7..2.........9..4.1#
23
1....6...4...8......9.2.5....4.9.3..6..5..9...7...1...........7..2.........9..4.1#
21
```

The puzzles dont come out with the same frequency.

Code:

```
   Number of Items: 11
             Total: 353.00
    Average (Mean): 32.09
            Median: 26.00
              Mode: 17.00
Standard Deviation: 16.01
```

Ive confirmed it even with a larger sample.

Maybe the effect will be very small for generating from a full grid though.

C

Last edited by coloin on Mon Jul 13, 2009 8:39 am; edited 1 time in total

Back to top profile pm

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, 4, 5 Next

Page 4 of 5

Stop watching this topic

Jump to: General/puzzle Go