Sudoku Players' Forums

? FAQ   **Q** Search   **≡** Memberlist   **□** Usergroups   **☑** Register

□ Profile   **△** Log in to check your private messages   **◎** Log in

THE REAL DISTRIBUTION OF MINIMAL PUZZLES

Goto page **Previous** **1**, **2**, **3** ... , **18**, 19, **20** **Next**

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

View previous topic :: View next topic

Author	Message
denis_berthier	□ Posted: Mon Sep 21, 2009 10:57 am Post subject: [quote]
	Eleven,
Joined: 19 Jun 2007 Posts: 828 Location: Paris, France	Thanks for the implementation of this optimisation.
	It now seems to be 5 times faster than before.
Back to top	[profile] [pm] [www]
gsf	□ Posted: Mon Sep 21, 2009 11:31 am Post subject: [quote]
	m_b_metcalf wrote:
Joined: 21 Sep 2005 Posts: 3811 Location: NJ USA	

> **gsf wrote:**
>
> > **m_b_metcalf wrote:**
> >
> > > **Code:**
> > >
> > > ```
> > > 1) generate a random complete grid
> > > 2) loop over all cells:
> > >     let P be the current puzzle
> > >     2a) choose one clue randomly
> > > from P and delete it, you get a
> > > puzzle P2
> > >     2b) if P2 has several
> > > solutions, restore the clue and GOTO
> > > 2a
> > >     2c) otherwise, set P=P2
> > >   end loop
> > > 3) return P (which is minimal).
> > > ```

Mike, two questions:

you have "2a) choose one clue randomly" inside a "loop over

> all cells" -- how does the loop affect the choice?
>
> how do you break out of the loop to get to "3)"?

Glenn,

I copied the term loop from Denis' description. It means looping over a rank-1 array containing the numbers 1 to 81 in random order. These are used to index the cells. The loop then just terminates. Probably better to say 'cycle' rather than 'GOTO 2a'.

thanks
I don't get what the cell selected by the outer loop index has to do with the randomly chosen clue in the inner loop
isn't "2a) choose one clue randomly" equivalent to "choose one cell randomly"?

Back to top   [profile] [pm] [www]

---

eleven

Joined: 10 Feb 2008
Posts: 483

Posted: Mon Sep 21, 2009 1:30 pm   Post subject: [quote]

> **denis_berthier wrote:**
>> Thanks for the implementation of this optimisation..

You are welcome, i like nice ideas, which improve a program with little implementation effort.

Of course now it could miss the first minimal 40 clue 😀

Back to top   [profile] [pm]

---

m_b_metcalf

Joined: 15 May 2006
Posts: 2358
Location: Berlin

Posted: Mon Sep 21, 2009 8:57 pm   Post subject: [quote]

> **gsf wrote:**
>> [I don't get what the cell selected by the outer loop index has to do with the randomly chosen clue in the inner loop
>> isn't "2a) choose one clue randomly" equivalent to "choose one cell randomly"?

Code:

```
1) generate a random complete grid
2) loop over all cells at random:
    let P be the current puzzle
    2a) delete the clue in the cell, you get a puzzle
P2
    2b) if P2 has several solutions, restore the clue
and cycle
    2c) otherwise, set P=P2
  end loop
3) return P (which is minimal).
```

Is this tauter (the wording, not the code -- that stays the same)?

Regards,

Mike

[Edit to repalce Obs. spelling by modern one.]

Last edited by m_b_metcalf on Tue Sep 22, 2009 12:27 pm; edited 1 time in total

Back to top   [profile] [pm]

denis_berthier   ◻ Posted: Mon Sep 21, 2009 10:06 pm   Post subject: [quote]

Joined: 19 Jun 2007
Posts: 828
Location: Paris, France

> **eleven wrote:**
> > Of course now it could miss the first minimal 40 clue 😃

Yes, but I now have almost two hundred thousand minimal puzzles generated without the optimisation and no 40, no 39, ..., no 32.
I think we can safely conclude that the chances of finding a 40 (or even a 39, ..., a 35) by this random generation are infinitesimal. Keeping the possibility of finding 32's, ..., 34's is already taking a very large safety margin.

I now have enough 30's to get a precise estimation of the complexity distribution of the 30's.
What I need now is as many 31's as possible in order to get a reasonable estimation of the mean complexity for the 31's.
Going beyond seems unrealistic, even after the optimisation. Moreover, I won't have much more time for this.

Back to top   [profile] [pm] [www]

m_b_metcalf   ◻ Posted: Mon Sep 21, 2009 11:51 pm   Post subject: [quote]

Joined: 15 May 2006
Posts: 2358
Location: Berlin

> **denis_berthier wrote:**
> > **m_b_metcalf wrote:**
> > > Before I look into this, can you tell me what the approximate yield is (minimal puzzles/million grids, say)?
> >
> > About 1 minimal puzzle for 225000 complete grids. On a standard 3 Ghz processor, this gives about 200 puzzles per day.

I now have a job running, starting with 35 randomly selected clues (i.e., dropping 46). It's producing a puzzle per min at 2GHz, one every 275,000 tries.

As you have so many already is there any further interest in this? I don't mind letting it run for a few hours, but not days.

Regards,

Mike Metcalf

Back to top [profile] [pm]

denis_berthier   Posted: Tue Sep 22, 2009 12:03 am   Post subject: [quote]

Joined: 19 Jun 2007
Posts: 828
Location: Paris, France

> **m_b_metcalf wrote:**
> I now have a job running, starting with 35 randomly selected clues
> (i.e., dropping 46). It's producing a puzzle per min at 2GHz, one every
> 275,000 tries.
> As you have so many already is there any further interest in this? I
> don't mind letting it run for a few hours, but not days.

The main interest I see here is checking the results with an independent
generator.
If you can produce a thousand puzzles, this should be enough.
Regards

Back to top [profile] [pm] [www]

m_b_metcalf   Posted: Tue Sep 22, 2009 12:24 am   Post subject: [quote]

Joined: 15 May 2006
Posts: 2358
Location: Berlin

> **denis_berthier wrote:**
> The main interest I see here is checking the results with an
> independent generator.
> If you can produce a thousand puzzles, this should be enough.
> Regards

OK. To whet your appetite, from the first 38 I obtain:

Code:

```
24    2
25   16
26   11
27    7
28    2
```

The SE ratings are:

Code:

```
2.0/1.2/1.2
8.3/1.2/1.2
3.4/1.2/1.2
2.0/1.2/1.2
7.2/1.2/1.2
1.5/1.2/1.2
6.6/1.2/1.2
2.0/1.2/1.2
2.0/1.2/1.2
7.2/1.2/1.2
2.0/1.2/1.2
2.0/1.2/1.2
6.6/1.2/1.2
7.3/1.2/1.2
2.0/1.2/1.2
7.3/1.2/1.2
2.0/1.2/1.2
6.6/1.2/1.2
```

```
7.2/1.2/1.2
2.0/1.2/1.2
6.6/1.2/1.2
1.7/1.2/1.2
3.6/1.2/1.2
6.6/1.2/1.2
2.6/1.2/1.2
3.4/1.2/1.2
7.1/1.2/1.2
7.2/1.2/1.2
7.3/1.2/1.2
1.5/1.2/1.2
1.2/1.2/1.2
7.2/1.2/1.2
2.0/1.2/1.2
1.5/1.2/1.2
7.2/1.2/1.2
7.2/1.2/1.2
1.5/1.2/1.2
4.2/1.2/1.2
```

Regards,

Mike Metcalf

Back to top   [profile] [pm]

denis_berthier	Posted: Tue Sep 22, 2009 12:34 am Post subject: [quote]

Joined: 19 Jun 2007
Posts: 828
Location: Paris, France

Mike,
In the SE ratings, I understand the first column but what are the 2nd and 3rd (all 1.2)?

Another variable of interest is the mean number of complete grids you need to generate a minimal puzzle.

Back to top   [profile] [pm] [www]

m_b_metcalf	Posted: Tue Sep 22, 2009 12:42 am Post subject: [quote]

Joined: 15 May 2006
Posts: 2358
Location: Berlin

> **denis_berthier wrote:**
>
> **Mike**,
> In the SE ratings, I understand the first column but what are the 2nd and 3rd (all 1.2)?
>
> Another variable of interest is the mean number of complete grids you need to generate a minimal puzzle.

That mean has moved to 350,000 now (?).

The second and third columns are useful only for the Patterns Game (first solution and first elimination). Ignore them.

Regards,

Mike Metcalf

Back to top   [profile] [pm]

denis_berthier   Posted: Tue Sep 22, 2009 12:45 am   Post subject: [quote]

Joined: 19 Jun 2007
Posts: 828
Location: Paris, France

> **m_b_metcalf wrote:**
>
> > **denis_berthier wrote:**
> >
> > > Another variable of interest is the mean number of complete grids you need to generate a minimal puzzle.
> >
> > That mean has moved to 350,000 now (?).

I think it isn't yet stabilised.
With suexg-cb, the fluctuations are very large.

Back to top   [profile] [pm] [www]

gsf   Posted: Tue Sep 22, 2009 8:12 am   Post subject: [quote]

Joined: 21 Sep 2005
Posts: 3811
Location: NJ USA

> **m_b_metcalf wrote:**
>
> > **gsf wrote:**
> >
> > > [I don't get what the cell selected by the outer loop index has to do with the randomly chosen clue in the inner loop
> > > isn't "2a) choose one clue randomly" equivalent to "choose one cell randomly"?
> >
> > **Code:**
> >
> > ```
> > 1) generate a random complete grid
> > 2) loop over all cells at random:
> >     let P be the current puzzle
> >     2a) delete the clue in the cell, you get a
> > puzzle P2
> >     2b) if P2 has several solutions, restore
> > the clue and cycle
> >     2c) otherwise, set P=P2
> >   end loop
> > 3) return P (which is minimal).
> > ```
> >
> > Is this taughter (the wording, not the code -- that stays the same)?

thanks
now I get it
and with a bonus vocab word

Back to top   [profile] [pm] [www]

m_b_metcalf   Posted: Tue Sep 22, 2009 9:34 pm   Post subject: [quote]

Joined: 15 May 2006
Posts: 2358

> **denis_berthier wrote:**
>
> If you can produce a thousand puzzles, this should be enough.

Posts: 2358
Location: Berlin

Denis,
I'm sending you the file by PM. Summary:

Code:

```
    953      25.7

     22        0
     23       13
     24       83
     25      307
     26      350
     27      164
     28       31
     29        5
     30        0
```

The rate was 1 puzzle per 287,000 tries, but note that when starting from only 35 clues many tries are 'still-born'.

The SE coverage is (ignore the second column):

Code:

```
1.2     31
1.5    109
1.7     22

2.0    251
2.3    765
2.5    664
2.6    845
2.8    872

3.0    408
3.2    595
3.4    793
3.6     67

4.2    553
4.4    837
4.5    117
4.6    914

5.6    678
5.7    600

6.2    413
6.6    740
6.7    491
6.8    818
6.9    118

7.0    707
7.1    392
7.2    402
7.3    777
7.4    127
7.5    643
7.6    748
```

```
7.7    570
7.8    770

8.2    545
8.3    193
8.4    380
8.5     60
8.9    482

9.0    619
```

Regards,

Mike Metcalf

Back to top   [profile] [pm]

denis_berthier   Posted: Wed Sep 23, 2009 9:31 am   Post subject: [quote]

Joined: 19 Jun 2007
Posts: 828
Location: Paris, France

Mike,
I didn't have much time today to study your results in detail.
Some aspects of these data and those from suexg-cb seem to be consistent:
- the mean (25.7) is within the suexg 95% confidence interval around 25.649
(with sd=1.122)
- the correlation coefficient #clues vs SER is 0.2, as for suexg-cb
- the unbiased average computed from your sample is 26.505 (from suexg-cb:
26.567)

But some aspects don't seem to be consistent:
- the mean value for the number of complete grids needed per puzzle (287000)
is much above the upper bound of the 95% confidence interval obtained from
suexg-cb: 224994 +- (1.96 * 224758 /sqrt(953)) ; it may be the case that your
complete grids have fewer minimals; Red Ed had noticed more minimals than
normal in suexg's generator part.

Last edited by denis_berthier on Thu Sep 24, 2009 3:29 am; edited 1 time in
total

Back to top   [profile] [pm] [www]

denis_berthier   Posted: Wed Sep 23, 2009 8:16 pm   Post subject: [quote]

Joined: 19 Jun 2007
Posts: 828
Location: Paris, France

Mike,
Do you apply any of the optimisations discussed in this thread (apart from
deleting the first 46 clues)?
A some point, you spoke of a random vector from which you choose the next cell
for deletion. Do you re-init it for each new complete grid?

Eleven, could you say a few words on how the next cell is chosen for deletion in
suexg. Is it completely random and in an order independent of that used for the
previous complete grids?

Last edited by denis_berthier on Thu Sep 24, 2009 3:30 am; edited 1 time in

total

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 - 8 Hours

Goto page **Previous** **1**, **2**, **3** … , **18**, 19, **20** **Next**

Page 19 of 20

Jump to: General/puzzle ⬦   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

Powered by phpBB © 2001, 2005 phpBB Group