# The hardest sudokus (new thread)



## Re: The hardest sudokus (new thread)

QUOTE

by **champagne** • Sun Jan 02, 2022 7:38 am

> **champagne** wrote:
> I was also surprised by the large difference in rating between SER 11.7 and sikh 11.0
> So I checked my earlier results and found that the skh had rated this puzzle as 11.5
> I have repeated the skh for both puzzles and this was the result:
>
> ```
> ....1....23...2.4...5...62.7.8.1B..7.5.56B.2721..6.7.B..5..ED=11.01.21.2
> 9E.76.5.7S....64.9..5.84.7..7.5E4.5.9B...9.....8..5....3.2 ED=11.51.21.2
> ```
>
> Quite different for essentially the same puzzle. But the gap between 11.5 and 11.7 for the canonical form is substantially lower than for the minlex form.
> To be sure that both puzzles are isomorphs of one another, I submitted the pair to my equivalencerelationship checker and here is the report.
> [hidden]Result of relationship analysis between ED=11.01.21.2 er ED=11.51.21.2

May be more about the story SER/SKFR

SER is known to be very very slow but is the referee for the pattern game. Nobody wanted to change this referee and to-day , AFAIK, no new rating was proposed leading to a common agreement to replace the well known rating given by SER(written in 2005??).

SKFR has been written to propose a kind of pre rating not too far from SER, not more.

SER is far from being bug free, this is true also for SKFR.SKFR has been written based on an old frame of my first solver. (full tagging). I saw quickly that the full tagging was far from the process applied by skilled players, so I switched to another frame and I use something similar in the view of SFR but much faster to-day when I am looking for a pre rating for SER.

This explains that SKFR will stay with identified bugs. It remains, in my view a tool for the gaming, nothing more.

The design of skh was done to approach the rating of SER giving usually, in the high ratings, something slightly smaller than SER.
As SER,_SKFR in the high rating can deliver different ratings for different morphs. If the process does not fall in a bug trap, the deviation remains relatively small +-.1 this is a well known situation in the pattern game.

The deviation can be much bigger if a UR rule is used in one run and not in the other (not exactly the same order to apply the uniqueness in both cases when several possibilities show up in the same time). This is normally of small influence in the pattern game, but leads normally to the biggest seen deviations SER versus SKFR..

SER has been reshaped recently, so is still "live". It would be simpler to keep skh as pre rating system and to focus on SER for discussions.

## Re: The hardest sudokus (new thread)

QUOTE

by **denis_berthier** • Sun Jan 02, 2022 8:23 am

> **champagne** wrote:
> It would be simpler to keep skh as pre rating system and to focus on SER for discussions.

I totally agree.

## Re: The hardest sudokus (new thread)

QUOTE

by **mith** • Mon Jan 03, 2022 4:28 pm

> **denis_berthier** wrote:
> However, I haven't checked all the rules in detail, but there may be more than rules for uniqueness without the confluence property.
> In particular, **as the thresholds between different ratings are based on totally arbitrary numbers of "nodes" (an extra-logical entity), it is very likely that the rules corresponding to a given rating don't have the confluence property. As a result, some resolution path with these rules may find a solution with SER = 11.0, while another path followed by a morph will not find one (and will require higher SER)**.

champagne may be able to better comment on this due to the work on skh, but my understanding of the SER rules is that - apart from uniqueness rules - the elimination of a candidate should never cause the minimum rating for the elimination of another candidate to rise. For chains in particular, the number of nodes for a given chain should never increase due to another chain being applied first. So I would think the confluence property should hold if uniqueness were removed.

That said, I don't know whether removing individual uniqueness techniques in the command line also disables them within dynamic chains; it may be something that needs to be disabled in the chain code separately.

(If I were to create a rating system from scratch, I would probably exclude uniqueness from it. There are certainly plenty of examples of puzzles that are very hard without uniqueness and trivial with, but I think I would prefer to live with that.)

## Re: The hardest sudokus (new thread)

QUOTE

by **mith** • Mon Jan 03, 2022 4:42 pm

(And I agree on the skf-as-pre-rating comments as well. The whole reason this discussion started is that I do use it as a pre-rating, and sometimes that causes 11.6+ SER puzzles to get shifted well down the priority list due to much lower skfr. q2 is likewise currently used as a pre-rating filter, due to the correlation between very high q2 and very high SER - it's certainly not perfect, but there will always be a trade off between speed and finding all puzzles in a neighborhood.

All that said, there may well be some correlation between SER-skfr delta and SER-SER morph deltas; if the former is being caused by different order of techniques, some of which may involve uniqueness, there may well be certain morphs which cause the same behavior in SER. Of course, it's possible - even likely - that there are some SER *and* skfr 11.0 puzzles which have morphs with significantly higher SER. Not much I can do about that other than rating a bunch of morphs though.)

## Re: The hardest sudokus (new thread)

QUOTE

by **champagne** • Mon Jan 03, 2022 4:41 pm

> **mith** wrote:
>
> champagne may be able to better comment on this due to the work on skh, but my understanding of the SER rules is that - apart from uniqueness rules - the elimination of a candidate should never cause the minimum rating for the elimination of another candidate to rise. For chains in particular, the number of nodes for a given chain should never increase due to another chain being applied first. So I would think the confluence property should hold if uniqueness were removed.

In SER, the basic rules should not give different ratings for different morphs. Counting the nodes is one way to qualify chain length limiting the effect of "sub chains redundancy". The source of small variations in ratings has 2 main reason (other are bugs):

The back search of the shortest length to explain a contradiction does not cover all possibilities ( explaining for example the difficulty to get a diamond 7.7)
For the highest ratings, to reduce the run time, the search does not cover all the field.

## Re: The hardest sudokus (new thread)

QUOTE

by **mith** • Mon Jan 03, 2022 4:58 pm

Ok, thanks champagne. So without those two speed improvements (and bugs) and uniqueness, it should be consistent. Perhaps with some faster implementation these shortcuts could be removed.

## Re: The hardest sudokus (new thread)

QUOTE

by **denis_berthier** • Tue Jan 04, 2022 7:00 am

> **mith** wrote:
>
> **denis_berthier** wrote:
> However, I haven't checked all the rules in detail, but there may be more than rules for uniqueness without the confluence property.
> In particular, **as the thresholds between different ratings are based on totally arbitrary numbers of "nodes" (an extra-logical entity), it is very likely that the rules corresponding to a given rating don't have the confluence property. As a result, some resolution path with these rules may find a solution with SER = 11.0, while another path followed by a morph will not find one (and will require higher SER)**.
>
> champagne may be able to better comment on this due to the work on skh, but my understanding of the SER rules is that - apart from uniqueness rules - the elimination of a candidate should never cause the minimum rating for the elimination of another candidate to rise. For chains in particular, the number of nodes for a given chain should never increase due to another chain being applied first. So I would think the confluence property should hold if uniqueness were removed.

I have never examined the SE code (except to notice that it is totally undocumented java code). However, what is clear from the output is, there are no real "chains" and no notion of chain length; there are inference networks with a certain number of nodes.
The underlying algorithm is T&E with some control on the number of inferences/nodes (the arbitrary thresholds). Would not it be for the thresholds (and disregarding uniqueness), I'd tend to agree that there is no problem of non-confluence in the "chains". But I can't see how SE could guarantee this in the presence of thresholds.

> **champagne** wrote:
> Counting the nodes is one way to qualify chain length limiting the effect of "sub chains redundancy". The source of small variations in ratings has 2 main reason (other are bugs):
> The back search of the shortest length to explain a contradiction does not cover all possibilities ( explaining for example the difficulty to get a diamond 7.7)
> For the highest ratings, to reduce the run time, the search does not cover all the field.

If this means anything, it is clearly not a convincing argument against the possibility of non-confluence when uniqueness is not involved.

> **mith** wrote:
> That said, I don't know whether removing individual uniqueness techniques in the command line also disables them within dynamic chains; it may be something that needs to be disabled in the chain code separately.

I wasn't aware that SE had uniqueness techniques within the dynamic chains. That's a really bad point. Are you sure of this?

> **mith** wrote:
> (If I were to create a rating system from scratch, I would probably exclude uniqueness from it. There are certainly plenty of examples of puzzles that are very hard without uniqueness and trivial with, but I think I would prefer to live with that.)

100% OK. And I wouldn't include any exotic pattern either.
As I've shown, adding rules for uniqueness (or exotic patterns) rarely changes the W rating and it's anyway interesting to know the cases when it does. For any rating system, that's possible only when such rules are not part of the fundamental ones.

## Re: The hardest sudokus (new thread)

QUOTE

by **champagne** • Tue Jan 04, 2022 7:53 am

> **mith** wrote:
>
> champagne may be able to better comment on this due to the work on skh, but my understanding of the SER rules is that - apart from uniqueness rules - the elimination of a candidate should never cause the minimum rating for the elimination of another candidate to rise. For chains in particular, the number of nodes for a given chain should never increase due to another chain being applied first. So I would think the confluence property should hold if uniqueness were removed.
>
> That said, I don't know whether removing individual uniqueness techniques in the command line also disables them within dynamic chains; it may be something that needs to be disabled in the chain code separately.

Hi Mith,

A little more on this.

No doubt, SER high ratings are **based on chains**. The node count is not more than a way to choose the chain (nest) to apply.
And you have no use of uniqueness rules in the chains.

Dynamic chains just expand a candidate using the "last in row,column, box"
Dynamic plus apply also patterns as locked sets or Xwings during the expansion phase
And nested chains apply as suggested in the name sub chains at each point of the expansion

## Re: The hardest sudokus (new thread)

QUOTE

by **denis_berthier** • Tue Jan 04, 2022 9:10 am

> **champagne** wrote:
> No doubt, SER high ratings are **based on chains**.

The typical meaningless claim, when you have no definition of a chain.
If you look at the SE output, they are clearly not chains (in the sense of continuous sequences of candidates).

> **champagne** wrote:
> And you have no use of uniqueness rules in the chains.

That's what I thought. Thanks for making it clear.

## Re: The hardest sudokus (new thread)

QUOTE

by **champagne** • Tue Jan 04, 2022 8:26 am

> **denis_berthier** wrote:
> No doubt, SER high ratings are **based on chains**.

The typical meaningless claim, when you have no definition of a chain.
If you look at the SE output, they are clearly not chains (in the sense of continuous sequences of candidates).

no comment

## Re: The hardest sudokus (new thread)

QUOTE

by **mith** • Tue Jan 04, 2022 6:37 pm

Thanks for the clarification, I had thought I had seen some uniqueness nested at some point, but I'm probably just conflating with yzf's UR chains.

## Re: The hardest sudokus (new thread)

QUOTE

by **mith** • Thu Jan 06, 2022 12:11 am

Apparently I had not already updated my local database to include hendrik's puzzles (nor JPF's additional puzzles found on that 11.8 pattern). I've done so now. I need to recheck the list jovi sent me a couple months ago, and otherwise I should be caught up other than the pattern game.

## Re: The hardest sudokus (new thread)

QUOTE

by **mith** • Thu Jan 06, 2022 5:09 am

A bunch of 11.7s, probably neighbors of some of those added puzzles:

```
CODE: SELECT ALL
.....1....23...4.15....6...7.187...9479...6.0.4...7.9.5....81.8..71.46  ED=11.7/1.2/1.2
.....1....23...4.15....6...7.187...9479...6.0.4...7.9.5....81.8..71.46  ED=11.7/1.2/1.2
.....1....2.34.25.....6.5T....7B9..8.6D..9.8D.78..8.6D...5.T2.9.....  ED=11.7/1.2/1.2
.....1....23...4.15....6...7.187...94.9...6.0.4....9.5T..81.8..71.46  ED=11.7/1.2/1.2
.....1....5.23....2585....6.5T....7B9.52..9.5B..9..6D.....96.T82..17.2.9..  ED=11.7/1.2/1.2
.....1....2.34.25.....6.5T....7B9.52..9.9.2B...1.9D.6.....6.7.5..8.6.....  ED=11.7/1.2/1.2
.....1....23...4.15....6...7.187...9479...6.0.4...7.9.5....81.8..71.46  ED=11.7/1.2/1.2
.....1....23...4.15.....6...7.187...4.7.8D..81.499.8..71..46...8.1.78..646....8.7  ED=11.7/1.2/1.2
.....1....2...3..45.....5.....6.437...8.9...2.9...1..4...56..6.7.3...  ED=11.7/11.7/3.4
```

## Re: The hardest sudokus (new thread)

QUOTE

by **hendrik_monard** • Thu Jan 06, 2022 1:56 pm

72 new 11.7's and a bunch of 11.6's (hidden)

```
CODE: SELECT ALL
98.7....7.68.....5..48..9..7..3...2....4....5..3..6..8.....7.6.....14.2.23;11.7;11.7;2.6
98.7B.5..7.5..B4..8..8.........9..5.....;32.6.94.7...5.8.....7.56..4.29;11.7;2.1;2.2
98.7B.5..7.498..645....4..95...976..5....87...6.....83....21.27;11.7;2.1;2.2
.....65.7.........6...43..72...3...6...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
98.7B.5..7.5..B4..8..8..87...794.8..B...95.....4.....8....4.5.....;32.27;11.7;2.1;2.2
98.7B.5..7.498..645....4..95...976..5....87...6.....83....21.27;11.7;2.1;2.2
98.7B.5..7.5..B4..8..8..4.B95....794.6.....87...6..8.....4.....;32.26;11.7;2.1;2.2
.....65.....76....4.732...6.41....12.43..34.76..2.31....16.27;11.7;2.1;2.2
98.7B.5..7.5..B4.....64..9...4.....33...8..5.6.94.7....7.56..4..98...;27.11.7;2.1;2.2
98.7B.5..7.498....68....5.95....4...87...974..8..4....6..8.....32.26;11.7;2.1;2.2
98.7B.5..7.498..645....4..95...976..5....87...6.....83....21.27;11.7;2.1;2.2
98.7B.5..7.5..B4..8..8..9..6..4..87...8..5B94..7....6.....9.3....7.26;11.7;2.1;2.2
.....65.....76....4.732...6.43.72...1..46...423.1..21.73....762.4...31..28;11.7;2.1;2.2
98.7B.5..7.498..645....4..95...976..5....87...6.....83....21.27;11.7;2.1;2.2
98.7B.5..7.498..645....4..95...976..5....87...79..6.8.4.....53....21.26;11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.67..2.74.1..1..2.;27.11.7;2.1;2.2
98.7B.5..7.5..B4.....64..9...4..87...12..3..34.67..2.74.1..1..2.;27.11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....26.1;27.11.7;2.1;2.2
98.7B.5..7.5..B4..8..8..9..6..4..87...974.6.5..95....6....5.....32.28;11.7;2.1;2.2
98.7B.5..7.498..645....4..95...976..5....87...79..6.8.4.....53....21.26;11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.67..2.74.1..1..2.;27.11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....26.1;27.11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....16.2;27.11.7;2.1;2.2
98.7B.5..7.498..4.5....6..95...4..87...79..6.8.4.....53....21.26;11.7;2.1;2.2
98.7B.5..7.4.5..6.4.89.6.8.7.9.5..9.6...4..8...;.....53....21.26;11.7;2.1;2.2
.....65.....6...43.72...1...46...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....16.2;27.11.7;2.1;2.2
98.7B.5..7.498..4.5....6..95...4..87...79..6.8.4.....53....21.26;11.7;2.1;2.2
98.7B.5..7.4.5..6.4.89.6.8.7.9.5..9.6...4..8...;.....53....21.26;11.7;2.1;2.2
.....65.....6...43.72...1...46...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....16.2;27.11.7;2.1;2.2
98.7B.5..7.498..4.5....6..95...4..87...79..6.8.4.....53....21.26;11.7;2.1;2.2
98.7B.5..7.4.5..6.4.89.6.8.7.9.5..9.6...4..8...;.....53....21.26;11.7;2.1;2.2
.....65.....6...43.72...1...46...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....26.1;27.11.7;2.1;2.2
98.7B.5..7.5..B4..8..8..9..6..4..87...974.6.5..95....6....5.....32.27;11.7;2.1;2.2
98.7B.5..7.5..B4..8..8..4..87...974.6.5..95....6....5.....32.27;11.7;2.1;2.2
.....65.....6...43.72...1...46...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
.....65.....6...43.72...1...46...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
98.7B.5..7.498..4.5....6..95...4..87...979.6.5..95....6....5.....32.27;11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....26.1;27.11.7;2.1;2.2
.....65.....6...43.72...1...46...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
98.7B.5..7.498..4.5....6..95...4..87...979.6.5..95....6....5.....32.27;11.7;2.1;2.2
.....65.....76....4.732...6..3...12.43..34.76..2..31....26.1;27.11.7;2.1;2.2
.....65.....6...43.72...1...46...423.1..4.1.73....674.2..3.1..;27.11.7;2.1;2.2
```

**Hidden Text: Show**

## Re: The hardest sudokus (new thread)

QUOTE

by **mith** • Thu Jan 06, 2022 2:48 pm

I had some of these in the database (mostly the 11.6s) but almost all not rated yet. One example of different SER for the morphs we are using in this batch:

```
CODE: SELECT ALL
.....1....2.3.2..4.5.....6..5..6..3.7289.....6..7..723.5..1...732.  ED=11.5/1.2/1.2
98.7..6..7..9.8...68..65..7.9.....5.6.4.....9..9...73...2.3.....1...9.  ED=11.6/1.2/1.2
```

(This one was actually in my March batch, but may have never been posted since it wasn't 11.6 on my side.)

I do have some ideas for catching more of the puzzles that my filters are skipping (without losing too much speed overall), will look into implementing that soon.



Powered by phpBB® Forum Software © phpBB Group
phpBB Mobile / SD | Privacy | Terms