

ヒントの少ない数独パズルの効率的な生成について

長尾 卓[†]
名城大学 理工学部[†]

山本 修身[‡]
名城大学 理工学部[‡]

1 はじめに

数独パズルは、ペンシルパズルの一種である。ペンシルパズルとは、問題に対して答えを鉛筆で書き込むパズルのことである。数独パズル以外にペンシルパズルは、スリザーリングや虫食い算などが知られている。数独パズルは、与えられたヒント (図 1) から、1 から 9 の数字を用いて縦、横、3 × 3 ブロックのどの数字も重複することのないように、マス埋めていくパズルである。図 1 で与えられる問題の答えは図 1 である。また、ヒントから導かれる最終盤面はただ 1 通りである必要がある。以降、数独パズルの答えを「解」と表す。

本研究は先行研究 [1] に基づいている。[1] は再急降下法を用いてヒントを生成しているが、シミュレーテッドアニーリングによって観測できなかったヒントの候補を考慮していなかった。そのため本研究ではその部分を考慮することで、効率的にヒントの生成をしようと考えた。本研究の目的は、[1] と同様に、ヒント数 17 の問題を効率よく生成することである。ヒント数の下限は 17 であることは [2] により証明されている。ヒント数 17 個の問題が難しい問題であるとは限らないが、本研究ではヒント数が最も少ない数独問題に注目する。

2 本研究で用いる手法

本研究では以下に説明するシミュレーテッドアニーリング (SA) とバックトラッキング (BT) の 2 つの手法を用いる。

SA は、ある状態からランダムな近傍状態に遷移させながら、局所最適解ではなく最適解を得るアルゴリズムである。内部温度というパラメータを設定し、時間が経つにつ

れて内部温度を低下させていき、その際、各状態の評価値が良くなる場合には確率 1 で遷移させ、悪くなる場合には温度に依存した確率で確率的に遷移させるのが一般的である。遷移確率としてメトロポリスのアルゴリズムを用いている。メトロポリスのアルゴリズムは、低評価への状態に遷移が可能であるため、理論上は局所最適解から脱することができ、最適解を得ることができる [3]。

メトロポリスのアルゴリズム [4] は、現状態を S_t 、遷移先の状態 S_{t+1} 、定常状態下でのある状態の確率 $P(S)$ 、 S_t から S_{t+1} へ遷移する確率を提案分布 $Q(S_{t+1}|S_t)$ 、負でない実数 r としたとき

$$Q(S_{t+1}|S_t)P(S_t)r = Q(S_t|S_{t+1})P(S_{t+1}) \quad (1)$$

と提案分布に対して、詳細釣り合い条件を成り立たすことができる。ただし、遷移確率 d は $d = \min(1, r)$ とする。また、提案分布に内部温度を示すパラメータを用いることができるため、SA との相性がよいことが知られている。

数独パズルを解く場合に、深さ優先探索 [5] (バックトラッキング: BT) を利用することができる。

3 ヒント生成のアルゴリズム

本研究では、最も解の個数が少なくなるようにヒントの候補を追加していく。本研究のヒント生成アルゴリズムの概略を図 3 に示す。SA によるヒント生成方法は、 n 個のヒントが生成されているとき、生成可能な解を SA を用いて等確率で 1,000 個生成し、最も出現頻度の少ないマスと数の組をヒントに追加する。先行研究 [1] では、現ヒントから生成可能な全ての解で、最も出現頻度の少ないマスと数の組と一致するという仮定の下行っていた。なお、本実験で SA の試行回数を増やすと解の個数はより減少するが、実行時間は長くなる。BT によるヒント生成方法は、BT を用いて現ヒントから生成可能な解を全て生成し、最も出現頻度の少ないマスと数の組をヒントに加える。なお、ヒント数 13 個以下は解の個数が膨大であり、全ての解を生成することは出来ないため、ヒント数 14 以上でこの方法を用いている。3 つまとめた BT とは、生成済みヒントを条件としたときに使用可能なマスと数の組を 3 つで 1 組とし、生成済みヒントに 1 組を追加した計 17 個の仮のヒントから生成できる解の個数を BT で調べ、最も解の個数が少なくなるような 1 組をヒントに追加する。この方法により、ヒント数 14 個の時点でヒント数 17 個の問題が生成可能であれば、必ずヒント数 17 個の問題を生成できることになる。なお、ヒント数 13 個から 4 つまとめて BT を行うと、計算量が膨大になるため、計算出来なかつ

			2	1				3
5	9			8				
	2	3		4				
1						5	9	
			6					
	3							4
			5				8	

7	4	8	5	2	1	9	6	3
5	9	6	7	8	3	2	4	1
3	2	1	9	4	6	8	7	5
6	5	2	3	9	4	7	1	8
1	3	4	2	7	8	5	9	6
9	8	7	6	1	5	4	3	2
2	7	3	8	6	9	1	5	4
4	6	9	1	5	2	3	8	7
8	1	5	4	3	7	6	2	9

図 1 17 個のヒントによる数独パズルの問題 (左) とその解 (右)

Efficient generation of Sudoku puzzles with few hints

[†] Suguru Nagao, Meijo University

[‡] Osami Yamamoto, Meijo University

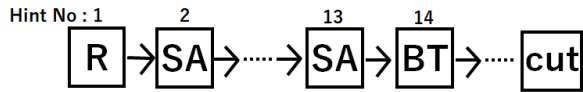


図2 先行研究のアルゴリズムの流れ [1]. 解の個数が1になった時点で処理 cut を行う. 処理 SA, BT, cut はそれぞれ SA による選択, BT による選択, マルコフ連鎖モンテカルロ法を用いてのヒント削減を表している.

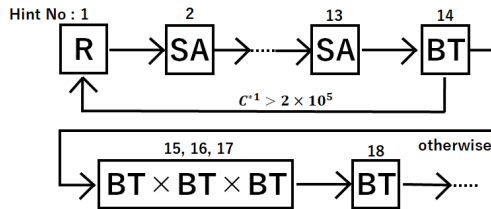


図3 提案するアルゴリズムの流れ. 解の個数が1になった時点で停止する. 処理 R はランダムな選択を表している. また C^*1 はヒント 14 が決定された時点での解の個数の見積である.

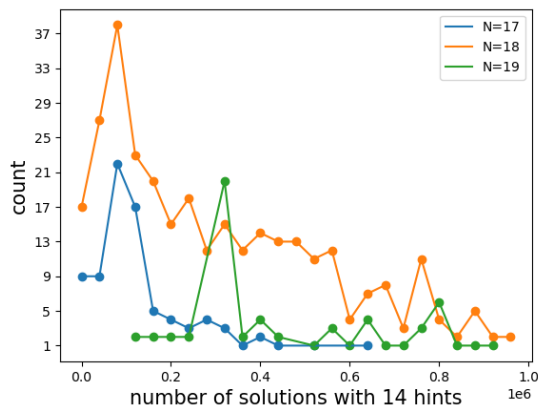


図4 最終ヒント数 (N) とヒント数 14 個時点での解の個数の関係

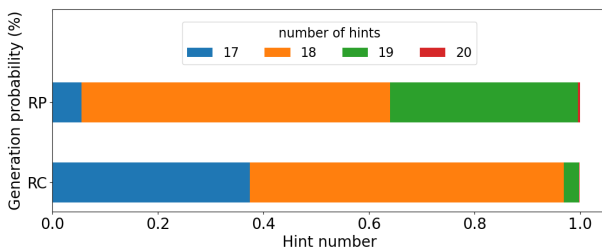


図5 先行研究 (RP) と本研究 (RC) ヒント数別の問題生成割合

た. ヒント数が 14 個の時点で解の個数が 2×10^5 個以下の場合に, ヒント数 17 個の問題が多く生成されていることが図 4 より確認できた.

4 研究結果

図 5 に先行研究と本研究のヒント数別の生成確率を示す. ヒント数 17 個の問題生成確率は, 先行研究では約 6%

であるのに対し, 本研究では約 37% に向上した. また, 問題を生成する平均時間は, 先行研究で約 1 時間 45 分であったのに対し, 本研究では約 1 時間 30 分と短縮することができた.

5 まとめと今後の課題

ヒント数 17 の問題生成確率が高い, ヒント数 14 個時点で解の個数が 2×10^5 個以下の場合のみ, ヒント 15~17 個を BT によってまとめて生成することを行った. その結果, ヒント数 17 個の問題生成確率が約 30% 向上した. また, ヒント数 14 個の時点で, ヒント数 17 個の問題生成が不可の場合があることと, 図 4 よりヒント数 14 個時点での解の個数は約 1.5×10^5 個を境に, ヒント数 17 個の問題生成確率は上がらないという傾向が確認できた.

また, SA によるヒント生成プロセスにおいて, SA に出現しないヒント (このヒントは急激に解集合を小さくする可能性が高い) を見逃さないで計算することが重要である. これは, 事前に BT で解の候補を取得して SA を行うことで解決することができるが, 現在はヒント数 12 個以上のときのみこの方法をとっている. 解の個数が 8 個以下の場合には計算量が膨大になってしまう問題点があり, 解の個数が 9~11 個の場合は, ヒントの組み合わせにより, 1 つのヒントの候補を調べる際に 10 分ほど BT を行うことがある. さらに今後の課題として, SA の試行回数を変化させヒント数 14 個時点での解の個数と実行時間のデータを取得し, 適切な SA 試行回数を調べることがある. 最後に, ヒント数 14 個目を複数試すことを行い, ヒント数 17 個の問題生成確率を高めていくことを行う. ヒント数 14 個時点での解の個数が 2×10^5 個の場合, 3 つまとめての BT の実行時間は約 10 分であるため, ヒント数 14 個目を 2~5 つほど試すことが可能と考えている.

参考文献

- [1] 古川 湧: ヒントの少ない数独パズルの生成に関する研究. 2020 年度名城大学大学院理工学研究科修士論文 (2021).
- [2] G. McGuire, B. Tugemann, and G. Civario: There is no 16-clue Sudoku: Solving the Sudoku minimum number of clues problem via hitting set enumeration. *Experimental Mathematics*, 23:2, pp. 190-217 (2014).
- [3] 池田 智悟, 窪田 耕明: シミュレーテッドアニーリング概説. http://mikilab.doshisha.ac.jp/dia/monthly/monthly00/20000415/ikeuchi_kubota.pdf, 参照 2021-01-06 (2000).
- [4] 野中良哲, 小野廣隆, 定兼邦彦, 山下雅史: メトロポリス・ヘイスティングスアルゴリズムに基づく $O(n^2)$ 到達時間ランダムウォーク. 平成 20 年度電気関係学会九州支部連合大会講演論文集, 04-1A-12 (2008).
- [5] 藤原 暁宏: アルゴリズムとデータ構造. 第 2 版, 情報工学レクチャーシリーズ, 森北出版 (2016).