

確率的サンプリングによるヒントの少ない数独問題の生成

古川 湧† 山本 修身‡

†名城大学大学院 理工学研究科 ‡名城大学 理工学部

1. はじめに

数独は、 9×9 マスから成るペンシルパズルの一種である。現在の数独は 1979 年ごろアメリカの建築家ハワード・ガーンズが発明したとされ [1]、ルールの簡単さから新聞やパズル雑誌など世界中で幅広く遊ばれている。数独の難易度は少ないほど難しい傾向にある。数独の問題をコンピュータに解かせる場合、バックトラックを用いることで難易度を問わず短時間で解くことができる。一方、問題を生成する場合は、解から数字を消していくという方法があるが、多くの場合、ヒント数が多い問題になる。本研究では確率的サンプリングと最急降下法を用いて、ヒントなしの状態からヒントを順次添加していくことでヒントの少ない数独問題を生成することを試みた。結果、図 1 に示すようなヒント数 22 の問題を生成した。

2. 提案する問題生成アルゴリズム

数独のヒント数の下限は 17 であることが知られている [2]。ここではヒント数 0 の状態からスタートして 1 ずつヒント数を増やし、ヒント数 17 の問題を生成することを試みる。ここで注目するのは与えられたヒントから得られる解集合の個数である。ある問題のヒントが $H_m =$ であるとする。これに対応する解を S_m とする。 $H_m = \phi$ のとき解集合は約 $|S_m| = 6.67 \times 10^{21}$ である [3]。このとき、 H_m にヒントを添加することでできる限り S_m を小さくすることを考える。あるヒントの集合 H_m を適正でない問題とする。解を複数求めることでヒント数 17 の問題が生成できる [4] が、本実験では解もしくは解に近い状態 S_m から、出現したマスと数字の組み合わせを求める。このとき最も出現回数の少なかった数字とマスの組み合わせを新たなヒントとして追加した集合を H_{m+1} とする。このような最急降下法を用いる。

3. Simulated Annealing を用いた探索

SA は、現在の状態から処理を行い次の状態に遷移したときに好ましい方向に向かった場合に温度変数 T を下げ最適解を見つけるアルゴリズムである [5]。本研究では問題の空白に適当な数字を入れ、それらの数字を解に近づくように 40 万回入れ替える。最初の 5 万回は雑音として評価せず、これ以降の位置と数字の組み合わせを状態とあわせて評価し、適切と思われるヒントを決定した。

3.1 メトロポリスのアルゴリズムとボルツマン因子

メトロポリスのアルゴリズムを用いることで評価の高い状態の出現割合を高くすることができる。状態空間 C から

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

図 1: 生成したヒント数 22 の問題。

評価値への関数 f が与えられているとき、定常確率分布が

$$P(C_i) = \frac{f(C_i)}{\sum_{C \in S} f(C)} \quad (1)$$

となるマルコフ連鎖を構成する。ただし、 S を全パターン集合とする。ある確率分布 $C(j|i)$ が与えられる。これに基づいて状態 C_i からその近傍状態 C_j へ遷移する確率と遷移せず留まる確率をそれぞれ

$$P(C_i \rightarrow j) = f(C_i)/f(C_j) \quad (2)$$

$$P(C_i \rightarrow i) = 1 - P(C_i \rightarrow j) \quad (3)$$

とすれば良い。 f は尤度を与えるもので、ここではボルツマン因子を用いて

$$f(C) = e^{-\beta E(C)} \quad (4)$$

と定義する。 β の値は SA の温度変数に相当しており $\beta = 1.5$ とした。 $E(C)$ は状態 C がどれほど解から遠いかを示す。 $E(C) = 0$ になれば、数の配置は解となる。

4. まとめと今後の課題

メトロポリスのアルゴリズムを用いたシミュレイトッドアニーリングを用いて数独の擬似解を列挙し、適切なヒントを添加する最急降下法を用いた結果、44 分かけて図 1 に示すようなヒント数 22 の問題を探索することができた¹。これは完全な解から問題を生成する方法 [4] のヒント数 17 の問題を 90 分かけて生成できたことに比べて、時間は短縮できたが大幅にヒントが増えている。今後の課題として、効率のよいヒントを探索できるように数値の調整を行う他、メトロポリスのアルゴリズム以外の手法を使うことによりヒントの少ない問題を生成する必要がある。今回生成した問題は中級者程度の問題だったが、人間が解くのが難しい問題の生成も課題である。

参考文献

- [1] Jean-Paul Delahaye: The Science behind SUDOKU. *Scientific American*, June, 2006, pp. 80-87 (2006)
- [2] G. McGuire, B. Tugemann, 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] B. Felgenhauer and F. Jarvis: Enumerating possible Sudoku grids. Technical Report pm1afj/sudoku/ (2005)
- [4] 古川湧, 山本修身: シミュレイトッドアニーリングを用いたヒント数 17 の数独問題の生成. WiNF2019, A3 (2019)
- [5] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi: Optimization by simulated annealing. *Science*, New Series, Vol. 220, No. 4598, pp. 671-680 (1983)

¹実行環境は OS: Linux, コンパイラ: gcc 5.4.0, CPU: Intel(R) Xeon(R) E5-2640 v4 @ 2.40GHz×2(40 スレッド)。