

数独パズルにおける初学者の熟達過程を 追跡するための問題生成と難易度評価の試み

Attempt to Generate Problems and Evaluate Difficulty to Track the Proficiency of Beginners in Sudoku Puzzles

尾関 智恵[†], 小笠原 秀美[‡]

Tomoe Ozeki, Hidemi Ogasawara

[†] 愛知工科大学, [‡] 中京大学

Aichi University of Technology, Chukyo University

ozeki-tomoe@aut.ac.jp

概要

数独はポピュラーなパズルであり、制約充足問題として知られている。これまで問題自体に対する提案は数多くあるが、人がどのように難易度を評価し、解答方略を獲得してパフォーマンスを向上させていく熟達過程に焦点を当てた研究は少ない。本発表では、数独における制約伝播に基づく解決方略を獲得していく過程の追跡をするため、これに適した問題生成とその難易度評価の探索的な試みを報告する。

キーワード：数独, 方略, 難易度, パフォーマンス

1. はじめに

数独¹[1] (ナンバープレース) は、古くから親しまれているパズルの一種で、代表的な制約充足問題である。このパズルは、 9×9 のセルを 3×3 のグループに区切り、縦・横・ 3×3 に1から9の数値をルールに従って当てはめていくペンシルゲームである。

このパズルの解答方略は、各セル目で候補となる値を探索的に確定させていく際、人は機械的方略のように網羅的に探索するわけではなく、部分的解答方法を用いた「解き筋」と呼ばれる概念(解き方の癖)が現れることがわかっている [2]。また、人にとっての問題の難易度判定の方法はさまざま提案されているが [3, 4, 5]、実際に解く時にかかる時間や認知的負荷も解き筋などの認知処理方略によって影響を受ける [6]。そのため、問題生成やヒントの出し方などの提案は数多くあるが、解答方略自体を経験的に変化させ、解く時にかかる時間や認知的負荷を軽減させるようなパフォーマンスを向上させていく熟達化に焦点を当てた研究は少ない。

本発表では、数独を題材に簡単な制約伝播に基づく解決方略を獲得していく過程の追跡をするため、これに適した問題生成と難易度評価の試みを報告する。

¹数独は株式会社ニコリの商標登録 (登 3225934) です。

2. 方法

2.1 熟達度合いを評価するための問題生成

ここでは人が問題を解いている過程のデータから、その解決者の方略の推定を容易にするような数独問題の生成について論じる。

2.1.1 制約伝播に基づく数独の解決方略

数独は制約充足問題であり、制約伝播に基づく多くの解決方略が見出されている [5]。本研究では以下の二つの数独の解決方略を対象とする。

方略 Single 数独ではあるセルと同じ行 (または列/ブロック) に属している他のセルには同じ数字は入らない。つまりあるセルに入りうる数字 (候補数字) は、同じ行 (または列/ブロック) に属している他のセルとは異なる数字である。もしあるセルの候補数字が一つしかない場合はそのセルにはその数字が入る (図 1)。この方略は Naked-Single とも呼ばれる。

方略 Hidden-Single 数独ではある行 (または列/ブロック) には数字 1 から 9 が一つずつ入る。もしある行 (または列/ブロック) に属するセル集合において、ある数字 n を候補数字とするセルが一つしかない場合はそのセルの値は n となる (図 2)。この 2 種類の手法は数独のルール説明に明記されている [1]。人による数独の解答過程はこの 2 種類の手法をそれぞれ必要とする度合いによって難易度が変化する可能性も示唆されている [2, 5]。

2.1.2 解決者の方略の推定方法

解決者が採用している方略の推定方法の一つは、解決者が記入したセルから推定することである。数独解

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

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

図1 Single

図2 Hidden-Single

決過程のある局面で解決者が記入したセルが特定の方略で決定可能ならば、解決者がその方略を採用している可能性がある。

もう一つの方法は問題を解くのに必要な方略から推定することである。ある方略 X の適用を繰り返すだけで解決できるが他の方略 Y の適用では解決できない問題を、解決者が解決できなかった(または解決に時間がかかった)とすればその解決者は方略 X をまだ獲得していないと推定できる。

第1の推定方法を効果的に適用するためには、その局面での空セルは対象となる方略(本研究では方略 Single または Hidden-Single)で決定可能であるが、どの方略でも決定可能なセルは少ないことが求められる。また両方略が適用可能なセルがバランスしていることも求められる。故に第1の方法で推定するためには以下の二つの指標が可能な限り小さい問題が求められる。なおここで S はある局面で方略 Single の一回の適用で決定可能な空セル集合を、 H は同様に方略 Hidden-Single の一回の適用で決定可能な空セル集合を、そして $n(S)$ は集合 S の要素数を示す。

$$\frac{n(S \cap H)}{n(S \cup H)} \quad \text{指標 1}$$

$$\frac{|n(S) - n(H)|}{n(S \cup H)} \quad \text{指標 2}$$

指標1はどちらの方略でも決定可能なセルの比率であり、これが0に近いほど方略推定できないケースが少なくなる。指標2は方略 Single と Hidden-Single で決定可能なセル数の差の比率であり、そのバランスの指標として用いられる。

第2の推定方法のためには、特定の方略のみを繰り返し適用することで解に到達できる問題が必要となる。今回は二つの方略を対象としているので問題は、それぞれの方略について、その方略のみで解に到達できる/できないの組合せで4つのクラスに分類される。表1に示すようにそれぞれの問題のクラスをそれぞれ sghsg, sg, hsg, ex と呼ぶ。第2の推定方法のためには

問題クラス sg または hsg の問題が必要となる。

表1 方略 Single/Hidden-Single と問題クラス

		Hidden-Single のみで	
		解決可能	解決不可能
Single のみで	解決可能	sghsg	sg
	解決不可能	hsg	ex

2.1.3 事前調査

上述した推定と難易度評価方法を観察的に検討するため、空セル数が40となる sghsg, sg, hsg の問題を生成し、2人の研究協力者による解答過程を記録した。その結果、sgの問題において解くための時間や認知的負荷が大きくなる、つまりパフォーマンスが下がる現象が観察された。具体的には解答完了までの時間が増大し、解答に4時間以上かかる場合もあった。hsgの問題については、1名は最後まで解答することができず断念していた。この他、hsg および sghsg の問題において、解答し終えたものの、候補の数値についてメモを取っていても選択を誤る場面が散見されている。

2.2 練習期間をはさんだ問題評価の実施

前節までを踏まえ実験用の問題を生成し、熟達度合いを評価できるかを検討した内容を以下に示す。

2.2.1 問題生成の方法

実験に用いた問題は Generate-and-Test によって生成した。つまりランダムに数独の解と空パターンを生成し問題を作成、その後一定の条件を満たしている問題を抽出した。なお事前調査の結果から、初級者でも極端に問題解決時間が長ならないように、本実験では空セル数30とした。また空セルのバランスを取るため各ブロックでの空セル数の差を1となるように空セルを設定した。ブロックごとの空セル数は3または4である。

第1の推定方法で解決過程の初期の段階を分析するために指標1、指標2を考慮して問題を生成した。具体的には次の2条件を設定した。

$$\frac{n(S \cap H)}{n(S \cup H)} \leq 0.1$$

$$\frac{|n(S) - n(H)|}{n(S \cup H)} \leq 0.5$$

ただし今回対象とする方略 Single と Hidden-Single について問題を生成する場合は方略 Single だけ決定可能なセルの数 ($n(S - S \cap H)$) は少なくなる傾向があった。その点を考慮し $\frac{n(S - S \cap H)}{n(S \cup H)} \geq 0.2$ という条件を追加した。この問題を Q1 と呼ぶ

第2の推定方法のためには問題クラス hsg (表1)の問題を用意した。この問題を Q2 と呼ぶ。なお第1の推定方法のためには問題 Q1 は問題クラス sghsg に属する。

2.2.2 実験協力者と倫理的配慮

愛知工科大学および愛知淑徳大学にて「数独に関する問題解決過程の調査」を目的とした研究であることを口頭・文章で説明し、同意した11名(平均年齢21.0歳)に実験協力者となってもらった。なお、本調査は愛知工科大学倫理審査委員会にて承認を得ている。

2.2.3 手順

実験協力者は愛知工科大学(5名)および愛知淑徳大学(6名)にて練習期間開始前日と最終日の次の日にそれぞれの大学で指定された小教室に集合した。

練習期間開始前日には、実験参加者に1人1台のタブレットとペンシル型入力機器が配布され、すべての作業をこれらの機器で実施するよう説明と練習の機会が与えられた。以後の活動は全て動画による画面キャプチャで記録するよう指示した。事前アンケートおよび練習問題3問(空セル数12で、ルールを理解しているかを確認)そして事前テストとして前節で生成した問題 Q1 と Q2 の2問を解いてもらった。

練習期間は4週間とし、1週間に7問ずつ解くようにタブレットにあらかじめ問題を用意した。この問題は空セル数27~40の問題が同程度にバランスするように組み合わせた計28問で、問題クラス sghsg, sg, hsg の3種類が同数になるように用意した。問題に取り組む時間やスケジュールは実験協力者の自由とした。

練習期間終了後の次の日には、事後テストを実施するために再度小教室に実験協力者が集合した。開始前日と全く同じ問題 Q1 と Q2 の2問を解いてもらった。その際、問題の難しさの感じ方に順番効果の影響が考えられたため、半数は問題順を逆にして取り組んでもらった。

2.2.4 実験デザインと評価方法

練習期間開始前と練習期間終了後の問題を解く時間に差が生じるかを確認するために、問題クラス(2水準:sghsg, hsg, 協力者内)×実施タイミング(事前・事後, 協力者内)の2要因の対応のある二元配置分散分析を行う。また、実験協力者が感じた問題に対する主観評価として、事前・事後テストでの問題解答直後に収集した「どちらの問題が難しいと感じたか」の解答結果を参考情報として報告する。

3. 結果

事前・事後テストでかかった解答時間(秒)を表2に示す。これらの解答時間について、対応のある二元配置分散分析を行なった結果、問題クラス間では hsg より sghsg の方が有意に時間が短く ($F(1, 40)=8.93, p<0.05$)、実施タイミング間では事前より事後の方が有意に時間が短かった ($F(1, 40)=4.15, p<0.05$)。解答時間の平均値の推移を図3に示す。

表2 事前・事後の解答にかかった時間(秒)

協力者	Q1: sghsg		Q2: hsg		難しい問題クラス
	事前	事後	事前	事後	
1	170	145	839	156	Q2: hsg
2	247	184	617	408	Q2: hsg
3	482	424	392	1359	Q2: hsg
4	347	280	1797	991	Q2: hsg
5	280	154	619	243	Q2: hsg
6	193	204	561	242	Q2: hsg
7	344	165	570	268	Q2: hsg
8	467	104	157	119	Q2: hsg
9	366	159	546	157	Q1: sghsg
10	389	158	175	157	Q1: sghsg
11	181	240	1247	369	Q2: hsg

事前・事後に収集した「どちらの問題が難しいと感じたか」の報告については、ほぼ全員が問題 Q2 (問題クラス hsg) を報告しているのに対し、2名が問題 Q1 (問題クラス sghsg) の方が難しいと一貫して報告していた。

4. 議論

熟達度合いを評価するために、制約伝搬に基づく数独の2種類の解決方略の使い分けが推定可能と考えられる問題クラスと空セル数の影響を想定した指標を整理した。事前調査を通して、熟達度合いを評価するた

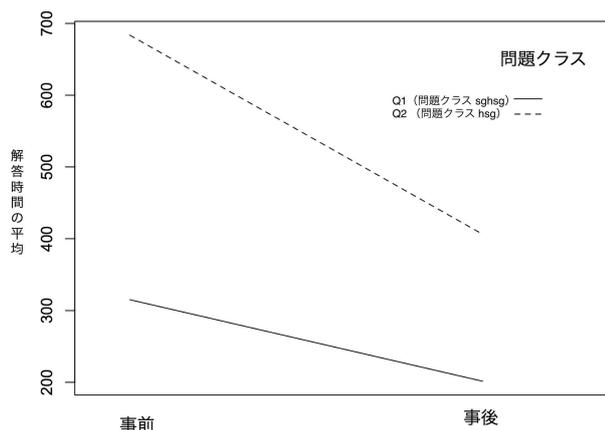


図3 問題クラスと実施タイミング別の平均値の推移

めには空セル数 30 で問題クラス sg hsg, hsg の問題が適切だと考え、これを使った実験調査を行った。

その結果、hsgの方が時間がかかっていたことから、Hidden-Singleのみで解く問題は認知的負荷がかかりやすい可能性がある。事前調査ではsgは4時間以上かかり、hsgは解くのを断念する協力者が見られたことから、SingleのみもしくはHidden-Singleのみで方略を変更できないことが解く時間を増大させる可能性がある。一方、sg hsgは解く時間もhsgより少なく、主観的申告でも多くが難しい問題に選ばなかったことから、方略を切り替えられることがパフォーマンスに影響する可能性が考えられる。これらの結果はSingleおよびHidden-Single 2種類の方略それぞれ必要とする度合いによって難易度が変化する報告 [2, 5] を支持する可能性がある。

今回の実験で用いた問題 Q1(問題クラス sg hsg) では記入セルで回答者が使用した戦略を判定できるようにしている。実際に解答過程を詳細に追跡することで熟達過程の分析に利用できるかを検討する必要がある。その上で得られた戦略の判定結果を問題クラス及び実施タイミングの解答時間の減少と主観的な難易度の報告と照らし合わせて解釈する必要がある。

5. まとめと今後の展開

現時点では熟達度合いを追跡する問題生成とその評価について、練習期間の事前・事後の実施結果の一部を整理した状態である。4週間の練習状況についても解答過程の画面キャプチャをおこなっており、今後はその記録からどのような過程を経てパフォーマンス向上つまり解答方略自体を経験的に変化させ、解く時にかかる時間や認知的負荷を軽減させるように変化したのかを分析する予定である。現段階ではコーディング

基準を設けて手作業で行っているが、画像解析で自動化する解析ツールの開発も並行して行っている。

また、熟達化過程の評価をどのように行うかについて、時間および主観以外の指標も検討する必要がある。熟達化は線形的に獲得されるものではなく [7, 8], 物理法則を発達段階でどのように獲得するかと同様にその過程を捉えるのは難しい課題である [10]。また、本研究で取り上げる熟達化過程の認知アーキテクチャ (Soar [9]) によるモデル化も検討する。

6. 謝辞

本調査に協力頂いた愛知淑徳大学人間情報学部の佐藤朝美教授及び加藤央昌准教授に謝意を表す。ご助言と自動解析ツールの開発に協力頂いている岐阜大学工学部の毛利哲也教授に謝意を表す。Soarによるシミュレーションを開発している中京大学工学部大石孝佑氏と手作業コーディングに従事している愛知工科大学工学部服部歩氏に謝意とメールを表す。

文献

- [1] 株式会社ニコリ, <https://www.nikoli.co.jp/ja/> (2023.4.11 閲覧)
- [2] 是川空, & 小谷善行. (2010). 数独の数理モデル「解き筋」. 第 51 回プログラミング・シンポジウム予稿集, 2010, 79-88.
- [3] 小場隆行, & 中所武司. (2011). 数独の難易度判定アプリケーションの提案と評価. 研究報告ゲーム情報学 (GI), 2011(8), 1-6.
- [4] 松原康夫. (2006). 数独の推論規則と難易度に関する考察. 情報処理学会研究報告エンタテインメントコンピューティング (EC), 2006(134 (2006-EC-005)), 1-6.
- [5] Pelánek, R. (2014). Difficulty rating of sudoku puzzles: An overview and evaluation. arXiv preprint arXiv:1403.7373.
- [6] Derbinsky, N., Bento, J., & Yedidia, J. S. (2014). Scalable methods to integrate task knowledge with the three-weight algorithm for hybrid cognitive processing via optimization. *Biologically Inspired Cognitive Architectures*, 8, 109-119.
- [7] 木村泉. (1998, June). 練習の中乘法則の折り紙実験による再検討. In 日本認知科学会大会発表論文集= Annual meeting of the Japanese Cognitive Science Society (Vol. 15, pp. 82-83).
- [8] 木村泉. (2003). 長期的技能習得データの「見晴し台」とその意義. 日本認知科学会 20 回大会発表論文集, 2003.
- [9] Karmiloff-Smith, A., & Inhelder, B. (1974). If you want to get ahead, get a theory. *Cognition*, 3(3), 195-212.
- [10] Laird, J. E. (2019). *The Soar cognitive architecture*. MIT press.