G-06

ナンバープレースの初期状態から難易度を判定できるか

岡澤翔太[†]，笠原祥平[†]，谷口大宙[†]

研究要旨

ナンバープレース（以下，ナンプレ）とは，９×９の計81マスの正方形の盤に，それぞれの行，列，ブロック（３×３の小さなまとまり）（以下，領域）に重複がないように１から９までの数字を埋めていくパズルである。ナンプレの問題を提供する企業，個人は数多く存在するが，それらの難易度を表す指標は出典ごとに様々で，統一のものが存在しない。そこで私たちは，ナンプレの難易度を推定する有効な共通の手段を作成することを目標に，問題と解答時間の関係について調べた。

Excelのマクロを用いて解答・作問・問題分析を行うプログラムを作成した。今回は，解法①（各マスが属する領域に存在する数字を調べ，候補から除外し，ある数字しか入り得ないマスがあればそれを確定する解法）と，解法②（解法①に加え，それぞれの領域ごとの候補を調べ，あるマスにしか入り得ない数字があればそれを確定する解法）を考えた。今回，ナンプレの問題を解くのに要した時間（以下，解答時間）を回答してもらうアンケートを実施した結果，私たちが作成した解法①で解ける問題と，解法①では解けないが解法②では解ける問題の間に解答時間の相関がないことが分かった。次に文献から引用した既存の問題を用意し，各問題について初期状態と，そこからすぐに得られる候補から得られるいくつかの情報を調べ数値化し，文献が提示している難易度や目標解答時間との相関を調べた。その結果，前述の数値同士にはそれぞれ線形的な相関がみられ，目標解答時間と各数値同士の関係は次の式で表されることが分かった。

$$T=\frac{q}{\log_{10} x - p}$$

このことから，調べた数値と目標解答時間をまとめたグラフの近似曲線の式を用いることで，任意の問題の目標解答時間，つまり難易度を推定できると考えた。

今後は，今回使用した問題とは異なる出典の問題についても考慮し，考案した難易度推定の手法がどれほど信用できるのかを検証していく必要がある。

１．はじめに

ナンバープレース(以下，ナンプレ)は９×９の四角い盤の中の各行，各列，３×３の各ブロックに１から９が１回ずつ入るように空いたマスを埋めるパズルである。はじめはマスがいくつか埋まっており，残りが空欄のままの盤の状態。上記のルールに合うようにマスを埋めていき，全てのマスが埋まれば完成である。パズルとして成立していれば，数字の入れ方は１通りしかない。この数学パズルは世界的に知られた有名なパズルであり，ナンプレの問題集やアプリ等では難易度別に問題が分けられている。しかし，それらの難易度を表す指標は出典ごとに様々で，統一のものが存在しない。

そこで私たちはナンプレの難易度を推定する有効な手段を作成することを目標に，問題と解答時間の関係を調べた。また解答する前に判定するため，初期状態から得られる情報のみで難易度を推定する方法について考えた。

今回は，ナンプレを解くのに要する時間が長い問題ほど，難易度が高い問題とみなした。

２．プログラムの作成

今回私たちは研究を進めていく中で２つの解法を考え，それらの解法を実装したプログラムを作成した。また，各解法で解くことが出来る問題の違いを調べるため，問題を作成するプログラムを作成した。

プログラムを作成していく中で，次の３つの盤を用意し，それらを用いてナンプレの状態を管理した。それぞれ，ナンプレの問題（初期状態）を入力する問題盤，問題を解いていく過程で，その時点で各マスに入りうる数字を記録しておく候補盤，実際に問題の解答を出力していく解答盤とした。

まず，ナンプレの問題を入力した後に，初めに候補の列挙を行った。この操作を初動操作と呼ぶこととする。ここで言う候補とは，ナンプレの基本ルール「各行，各列，各ブロックに重複がないように１から９の数字を埋める」に反しない数字とする。

初動操作の手順は次のとおりである。初めに候補として１から９までの数字を全て入力した。次に各マスを調べ，そのマスが属している行，列，ブロックにある数字を候補から除いた。全てのマスを調べた後に残った数字を，最終的な候補とした。（図１）

次に解法①（各マスが属する領域に存在する数字を調べ，候補から除外し，ある数字しか入り得ないマスがあればそれを確定する解法）では初動操作を行った後，候補盤の全てのマスを調べ，候補が１つしかないマスがあれば，その候補を解として解答した。その後，初動操作と同じような操作を経て候補盤を更新した。以上の操作を繰り返し，解答盤を更新することが出来なくなれば，それ以上解き進めることが出来ないと判断し，プログラムを終了した。（図２）

解法②（解法①に加え，それぞれの領域ごとの候補を調べ，あるマスにしか入り得ない数字があればそれを確定する解法）では解法①で候補が１つのみのマスを調べる直前に，各行，各列，各ブロックを調べ，１つのマスにのみ存在する数字があればその数字を解として解答する，という操作を追加した。解法①と同様に，解答盤が更新されなくなればプログラムを終了した。（図３）

また各解法で解くことが出来る問題を作るために，次のように作問した。まず81マス全て基本ルールに反しない形で埋められた状態（完成盤とする）を用意した。その後，ランダムにマスを選んで数字を消し，解法①か②を用いて完答出来るかどうか調べた。完答出来なかった場合は消した数字をもとのマスに戻した。10回連続で完答出来ない問題を作成した時に，プログラムを終了した。（図４）

† 兵庫県立宝塚北高等学校

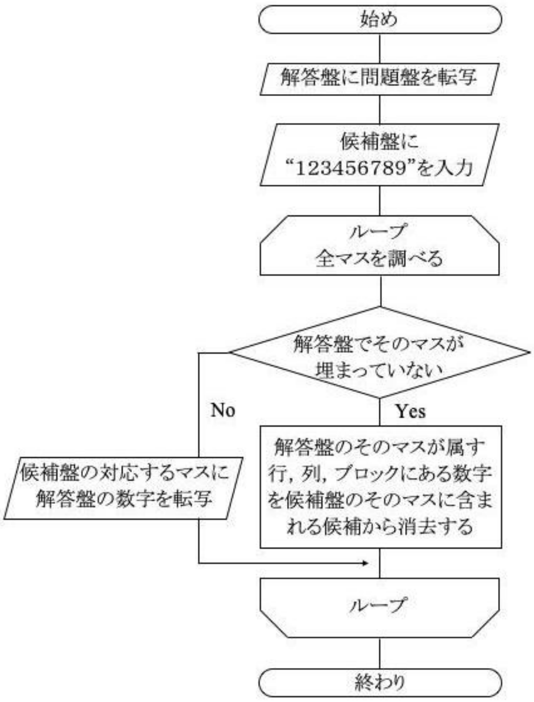

図１　初動操作のフローチャート

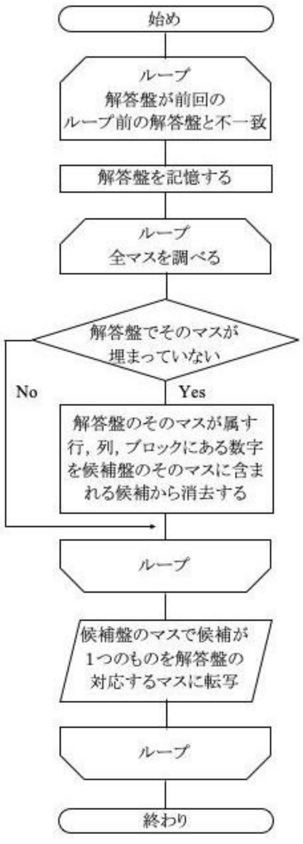

図２　解法①のフローチャート

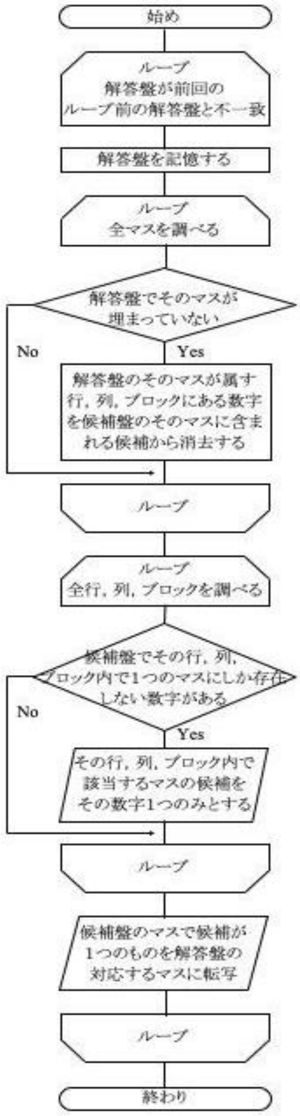

図３　解法②のフローチャート

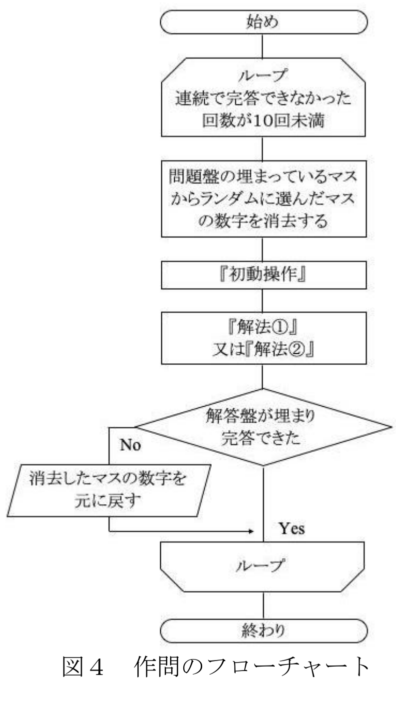

図４　作問のフローチャート

３．難易度と解法の関係

プログラムの構成上，解法①で完答出来る問題(以下，問題群①)は解法②でも完答出来る。しかし，解法②で完答出来るが解法①で完答出来ない問題(以下，問題群②)を複数確認したので，この逆は言えない。２種類の完成盤について，問題群①の問題と，問題群②の問題の計４種類の問題を作成した。その後，用意した４種類の問題を解くのに要した実際の時間を本校グローバルサイエンス科の有志（14名）に協力してもらい簡単な調査を実施した。解法②が解法①を含んでいるので，問題群②の問題の方が問題群①の問題よりも難易度が高いと考えられたが，アンケートの結果からはその傾向が見られなかった。このことから，完答できる解法によって難易度を区別することは出来ないと考えられる。(図５)

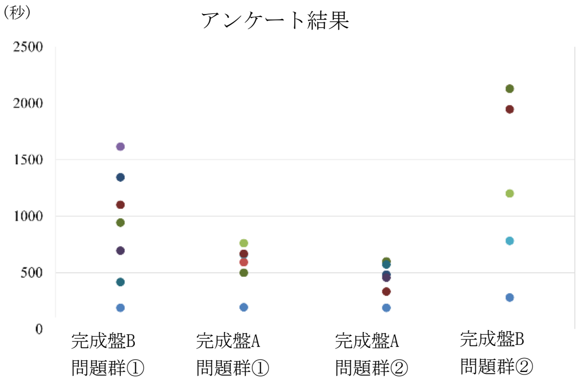

図５　各所要時間

解法②が解法①を含んでいるので，問題群②の問題の方が問題群①の問題よりも難易度が高いと考えられる。しかし，アンケートの結果からはその傾向が見られなかった。このことから，完答できる解法によって難易度を区別することは出来ないと考えられる。そこで，文献(4)から引用した69問について，解法①と解法②でそれぞれ解答できるかどうかを調べた。その結果，文献(4)が提示している目標解答時間が５分で解法①で完答できない問題や，目標解答時間が30分で解法①で完答できる問題が確認できた。このことから，完答の可否によって明確に難易度を区別することは出来ないと考えられる。（図６）

min	5	5	5	5	5	7	7	7	7	7	10	10	10	10	10	12	12	12	12	12	12	15	15	15	15	15	15	18	18	18	18	18	18	18
1	O	O	O	O	×	O	O	O	O	×	O	×	O	×	O	O	O	O	O	O	O	O	O	O	O	×	×	O	O	O	×	×	×	×
2	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O

min	20	20	20	20	20	20	20	23	23	23	23	23	23	23	23	25	25	25	25	25	25	25	25	25	25	30	30	30	30	30	30	30	30	30	30
1	O	×	O	O	×	O	×	O	O	O	×	O	O	×	×	O	×	×	×	×	×	×	O	×	×	O	×	×	×	×	×	×	×	O	×
2	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O

図６　目標解答時間と完答の可否

４．難易度と初期状態から得られる情報の関係

ナンプレの初期状態から得られる情報として，以下の６つのパラメータを設定した。以下，各行，各列，各ブロックをまとめて各領域と表現する。

A. その時点で埋まっているマスの総数
B. 各領域内の埋まっているマスの数に１を足したものの積
C. １から９について，それぞれの数字がその時点で埋まっている個数に１を足したものの積
D. 各マスの候補数の積
E. 各領域内の候補数の積
F. １から９について，それぞれの数字が入りうるマスの個数の積

これらのパラメータを用いた難易度推定の式を考案するため，まずは文献(3)(4)にある問題について各パラメータ同士の関係について調べた。前述の６つのパラメータのうちのどの２つについても線形的な相関が確認できた。このことから，難易度推定の式に前述のパラメータを２つ以上同時に使う必要はないと考えられる。（図７，８）

パラメータ		R^2
A	B	0.9929
A	C	0.9905
A	D	0.9733
A	E	0.9716
A	F	0.9758
B	C	0.9828
B	D	0.9722
B	E	0.9707
B	F	0.9795
C	D	0.9793
C	E	0.9809
C	F	0.972
D	E	0.9987
D	F	0.9931
E	F	0.9926

図７　文献(3)に基づく各パラメータ同士の関係

パラメータ		R^2
A	B	0.9372
A	C	0.917
A	D	0.9523
A	E	0.9443
A	F	0.93
B	C	0.8693
B	D	0.9065
B	E	0.897
B	F	0.9352
C	D	0.9583
C	E	0.9558
C	F	0.8732
D	E	0.9982
D	F	0.9548
E	F	0.9531

図８　文献(4)に基づく各パラメータ同士の関係

そこで，前述の６つのパラメータと，文献(4)が提示している目標解答時間との関係を調べた。その結果６つのパラメータと，文献(4)が提示している目標解答時間との関係は，次の式で高い精度で近似できた。（図9，10）

$$T=\frac{q}{\log_{10}x - p}$$

T:目標解答時間
x:６つのパラメータのうちの１つ
p，q:使用したパラメータによって定まる定数

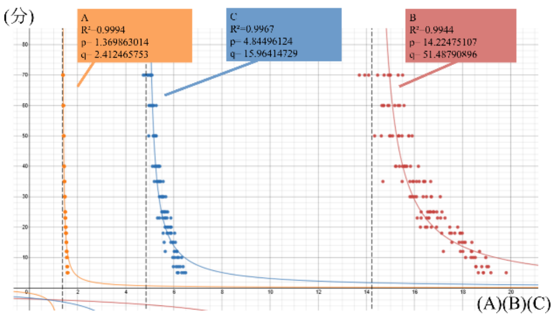

図９　パラメータA～Cと目標解答時間

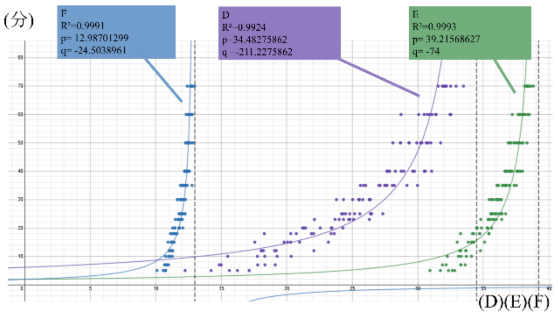

図10　パラメータD～Fと目標解答時間

このことから，前述の６つのパラメータのうち，どれか１つを調べ上記の式に代入することで，その問題の解答時間を推定できると考えられる。そこで考案した式を用いて文献(3)から引用した問題について前述の６つのパラメータを調べ，解答時間を推定する式に代入することで，実際に解答時間を推定した。今回，文献(4)に基づいて求めた２種類の式を使用した。

分		(パラメーター) A	B	C	D	E	F
	①	6.654	6.105	5.899	6.684	6.081	6.154
(目標回答時間)	②	53.479	49.565	59.646	59.668	59.687	48.978
	③	17.832	16.961	16.250	17.492	17.204	17.790

図11　文献(4)に基づく式を用いて推定された時間

図11より，文献(4)に基づいて得られた式をもとに解答時間を推定すると，どのパラメータを用いた時もおおよそ同じ値を示したため，いずれのパラメータであっても解答時間を推定できていると考えた。そこで文献(4)の問題について，各パラメータによって推定した時間と，各文献が提示している目標解答時間との誤差を評価するため，各問題について推定された解答時間と，文献が提示している目標解答時間との差の絶対値をとり，目標解答時間で割った値を調べた。（図12）最終的に得られた値が目標解答時間１分あたりの誤差であるので，これらの値の平均や分散が最も小さくなるパラメータＥ(各領域内の候補数の積)が，解答時間の推定に最も適切なパラメータであると考えられる。

パラメータ	A	B	C	D	E	F
平均	0.293	0.455	0.375	0.265	0.256	0.469
分散	0.209	0.995	0.661	0.100	0.0778	3.95

図12　文献(4)における１分あたりのずれ

５．結論と今後の展望

与えられた問題について，初期状態から得られる以下のパラメータを調べ，以下の式に代入することで，その問題の解答時間が推定できることがある。特に各領域内の候補数の積を用い，下記の式を用いると，その他のパラメータより小さなずれで解答時間を推定できると考えられる。

$$T=\frac{7.4}{\log_{10}E - 39.2}$$

T:目標解答時間
E:各領域内の候補数の積

この式を用いれば，例えば出典が異なる２つの問題の難易度を比較・推定することができる。このことから，今回考案した方法は，ナンプレの難易度を推定する有効な手法であると考えられる。

今回考案した難易度を推定する式によって得られた解答時間がどれほど信用できる値なのかを，様々な出版社，出題者の問題集を用いて検討するとともにアンケートなどによって調査できる実際の解答時間と比較することによって評価する必要がある。また，実際は解答者の熟練度やひらめきによりナンプレの解答時間が大きく変わるので，ナンプレの解法に関する定跡との関係性についても今後検証していきたい。

６．参考文献

(1)ジェイソン・ローゼンハウス,ローラ・タールマン 著，小野木明恵 訳(2014)『「数独」を数学する 世界中を魅了するパズルの奥深い世界』青土社
(2)Agnes M. Herzberg, M. Ram Murty (2007)”Sudoku Squares and Chromatic Polynomials” 「Notices of the AMS」 Vol.54 p.708-717
(3)Katsutoshi Seki (2018)「解独 - 数独・ナンプレの解析プログラム」https://sekika.github.io/kaidoku/ja/ (2019.11.1 閲覧)
(4)実力検定ナンプレ250問 2012年５月号．東京，株式会社コスミック出版，2012，p258.
(5)Carlos F. Daganzo (2018) “MINUET: A METHOD TO SOLVE SUDOKU PUZZLES BY HAND” arXiv」 https://arxiv.org/pdf/1812.06778.pdf
(6)Peter Norvig / 青木靖 訳 「あらゆる数独パズルを解く」http://www.aoky.net/articles/peter_norvig/sudoku.htm (2019.11.1 閲覧)