Vol.2017-GI-37 No.9
2017/3/7

パズルゲームの個人対応難易度評価

濱田 信佑[1,a)]　西野 順二[1]

概要：パズルゲームにおいて、ユーザーが未解答の問題に感じる難易度評価を、個人差に適応して推定する方法について検討した。これまで難易度評価は問題に固定のものであり、個人が感じる差異については扱われず、違和感を感じさせることがあった。パズル問題の解法特性を、発見的な解法アルゴリズムの適用可能性やその頻度を特徴量として用い、ユーザごとに実際に解いて評価した難易度との関係モデルを実験的に分析した。代表的なパズルゲームのお絵かきロジックの問題を対象とし、11 種の特徴量を設定した。難易度が異なると予想される 5 問を解いて、難易度の相対評価を訊ねる被験者実験を行った。約半数の被験者は回答にかかる推定手数と難易度に強い相関があった。一部には、ある種のアルゴリズムの使用可否や、推定手数のばらつきと難易度が相関するものもあり、個人によって問題の難易度を感じる原因が異なることが明らかになった。

1. はじめに

パズルゲームでは難易度を問題選択の指標とすることが多い。しかしパズルを扱う雑誌やアプリケーションなどの問題に記された難易度はユーザーの感じる難易度に一致するとは限らない。本研究の目的は個人の感覚に一致した難易度を提供する手法を提案することである。

パズルゲームであるお絵かきロジックについて、回答者が持っている難易度を決定するための要素に依存して難易度を評価する。お絵かきロジックに類似するパズルゲームである数独では、問題は解き方によっても難易度の感じ方が違うことが知られており [1] お絵かきロジックにおいても解き方の違いが難易度の感じ方の違いを生じさせていることが見込める。既存のお絵かきロジックの難易度評価 [2]、あるいは数独の難易度評価 [3][4] では個人ごとの難易度の基準の違いは考慮されていない。

図 1.1 の様なシステムを考える。お絵かきロジックの問題の盤面 p が与えられた時、p についての特徴量のベクトル $\boldsymbol{f}$ を抽出する。システムにはその $\boldsymbol{f}$ が入力されシステムはユーザーの難易度の感じ方の違いに対応して推定された難易度を出力する。

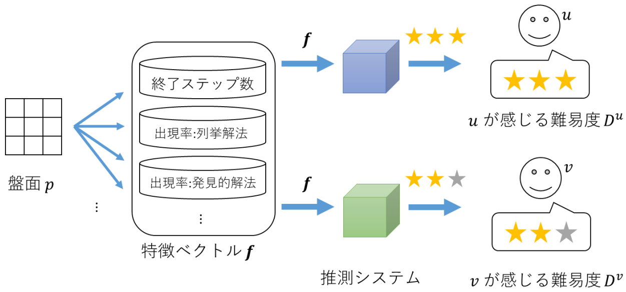
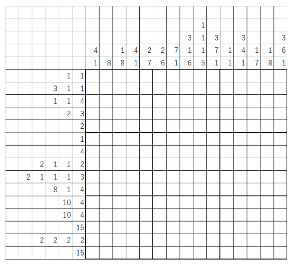


図 1.1　個人に対応する難易度推定システム

2. お絵かきロジック

お絵かきロジックとは与えられた縦・横のヒントの数字に基づいてマスを塗りつぶすことで絵を完成させるペンシルパズルである。ピクロス、ののぐらむ [5]、イラストロジック [6]、ペイントロジック [7] などの別名がある。



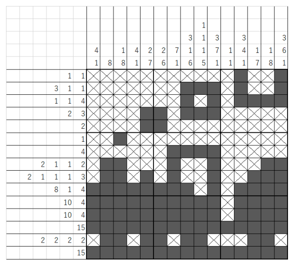


図 2.1　初期盤面 (問題 24446[8])**図 2.2**　完成盤面 (解答：「汽車」)

2.1　お絵かきロジックのルール

お絵かきロジックの開始時に初期盤面と縦向き・横向きのヒントの数字の組が与えられる。ヒントの数字に対して、完成盤面は必ず 2 つの条件を満たす。

- ヒントの数字とその順番がその行・列にある塗るマスの連続数と順番になる
- 塗るマスの連続の間には 1 マス以上の塗らないマスが

[1] 電気通信大学
[a)] h1311155@edu.cc.uec.ac.jp

ある
条件を逸脱しないようにヒントの数字から塗るマス・塗らないマスを推測して盤面を更新し、更新した盤面から推測することを繰り返して初期盤面から完成盤面へ辿り着く。まず図 2.1 の様に初期盤面が問題として与えられ、いくらかの更新を経て図 2.2 の様な完成盤面へと辿り着くことができれば正解である。

2.2 候補の列挙解法

盤面を更新する場合、たいていは 1 つの行・列に注目して更新する。更新はヒントの数字による束縛と行・列の更新前に確定していたマスの両方を逸脱しない全ての可能な塗り方、すなわち解候補を列挙し、その候補全てにおいて、あるマスが共通して塗る（塗らない）マスであるならば、そのマスを塗る（塗らない）マスとして盤面に記録する。この考え方が盤面更新の基本であり、これを列挙解法 E と呼ぶことにする。

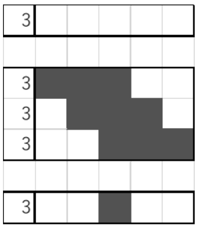


図 **2.3** 列挙解法による更新

人がお絵かきロジックを解く時には候補の全列挙をするとは限らず、テクニックとして様々な発見的な解法が用いられる。発見的な解法は列挙解法が適用できる行・列の内のいくつかで用いることができる。発見的な解法を用いるとより単純な思考で盤面を更新することができ、難易度の感じ方に影響を与えると考えられる。

3. 問題の特徴量抽出

本研究では、7 種類の解法を取り上げ、その出現率それぞれの平均 $\overline{d}(E), \overline{d}(h_1), ..., \overline{d}(h_6)$ および無作為更新の試行の終了ステップ数の平均 $\overline{T^E}$、分散 $\sigma^2(T^E)$、最小値 $T^E_{\min}$、最大値 $T^E_{\max}$ の合計 11 個の値を特徴量をとして提案した。以下では、各特徴量について述べる。

3.1 お絵かきロジックの盤面状態

1 つのマスは 3 種のうちのいずれかの状態として扱う。

- 確定した塗るマス
- 確定した塗らないマス
- 塗るか塗らないか未確定なマス

初期盤面では全てのマスは未確定なマスである。

お絵かきロジックの問題におけるある盤面を p とする。p は縦横の大きさを持ち、$n \times m$（縦 n マス、横 m マス）の大きさである時、p は n 個の行と m 個の列で構成される。

盤面 p の i 番目の行は $r_i(p)$、j 番目の列は $c_j(p)$ と表す。行と列を区別しなくてよい場合の 1 つの行・列を l と表す。

3.2 解法適用による更新のステップ数

1 つの行・列に注目するお絵かきロジックの解法を s とする。s を適用してある行・列 l はいくつかの未確定マスを塗るマスか塗らないマスとして確定することができる。解法 s を l に適用してできる行・列を $\mathrm{apply}(s, l)$ と表す。

任意の解法で更新可能な行・列の内の 1 つをある解法に基づいて更新することを 1 ステップと数えて、初期盤面からの最初の更新をステップ $t = 1$ とする。また、初期盤面から数えた完成盤面までのステップ数 T を終了ステップ数とする。

列挙解法 E で更新可能な行・列から 1 つを無作為に選んで更新することを繰り返した場合の終了ステップ数を列挙解法による無作為更新の終了ステップ数 T^E とする。また実際に初期盤面から完成盤面まで更新によって T^E を求めることを無作為更新の試行と呼ぶ。

T^E は無作為更新の試行ごとに異なる値を取り得る。n 回目の無作為更新の試行の終了ステップ数を T^E_n と表す。

3.3 解法出現率

t 回更新した $n \times m$ 盤面 p^t においてある 1 つの行・列に注目する解法 s で更新可能な行・列の数の総和を $U(p^t, s)$ とし、式 3.2 で定義する。

$$\mathrm{updatable}(s, l) = \begin{cases} 1 & \text{if } l \neq \mathrm{apply}(s, l) \\ 0 & \text{if } l = \mathrm{apply}(s, l) \end{cases} \tag{3.1}$$

$$\begin{aligned} U(p^t, s) = & \sum_{i=1}^{n} \mathrm{updatable}(s, r_i(p^t)) \\ & + \sum_{j=1}^{m} \mathrm{updatable}(s, c_j(p^t)) \end{aligned} \tag{3.2}$$

ある $n \times m$ 盤面 p についての解法 s の出現率 $d(s)$ を式 3.3 で定義する。

$$d(s) = \frac{1}{T^E} \sum_{t=0}^{T^E - 1} \frac{U(p^t, s)}{n + m} \tag{3.3}$$

$U(p^t, s)$ は t までに更新した行・列の更新順序に依存するため、$d(s)$ は無作為更新の試行ごとに異なる値を取り得る。n 回目の無作為更新の試行の解法 s の出現率を $d_n(s)$ と表す。

3.4 対象とした発見的解法

特徴量抽出の対象として 6 つの発見的解法を設定した。

h_1： 総和一致

h_2： 全ヒント確定

h_3： 両端の塗るマス注目

h_4： 両端の塗らないマス注目

h_5： 最大ヒント注目

h_6： 最小ヒント注目

これらの発見的解法はトリミングを併用しながら用いる。発見的解法の詳細とトリミングについては付録に記す。

3.5 抽出する特徴量

N 回の無作為更新の試行で得られた値の代表値を式 3.4〜3.8 で定義し、11 個の成分からなる特徴量のベクトル $\boldsymbol{f}(\mathrm{p})$ を式 3.9 で定義する。$\boldsymbol{f}(p)$ は初期盤面である盤面 p に N 回の無作為更新の試行をして得る。

$$\overline{d}(s) = \frac{1}{N}\sum_{n=1}^{N} d_n(s) \tag{3.4}$$

$$\overline{T^E} = \frac{1}{N}\sum_{n=1}^{N} T_n^E \tag{3.5}$$

$$\sigma^2(T^E) = \frac{1}{N}\sum_{n=1}^{N} (T_n^E - \overline{T^E})^2 \tag{3.6}$$

$$T_{\min}^E = \min_{n\in[1,N]} T_n^E \tag{3.7}$$

$$T_{\max}^E = \max_{n\in[1,N]} T_n^E \tag{3.8}$$

$$\begin{aligned}\boldsymbol{f}(p) = (&\overline{d}(E), \overline{d}(h_1), ..., \overline{d}(h_6),\\ &\overline{T^E}, \sigma^2(T^E), T_{\min}^E, T_{\max}^E)\end{aligned} \tag{3.9}$$

4. 個人特徴のモデル化

人ごとに異なる難易度の感じ方を測るため評価データの収集実験を行った。またデータ収集実験に用いた問題から特徴量を抽出した。

4.1 実験形式

お絵かきロジックの問題を 5 つ用意し、被験者はそれらの問題を解いた。被験者ごとに次の項目を記録した。

- それぞれの問題の解答に要した時間
- それぞれの問題の比較評価
- お絵かきロジックの経験の有無

被験者は 5 つの問題から 2 つを取り出す全ての組（${}_5C_2 = 10$ 通り）について 5 段階の比較評価を行った。2 つの問題を X, Y とすると以下の評価を行う。

(1) X の方がとても難しい
(2) X の方が難しい
(3) X と Y は同程度に難しい
(4) Y の方が難しい
(5) Y の方がとても難しい

4.2 使用問題

データ収集実験で使用した問題とその大きさを表 4.1 に記す。問題番号とは引用元 [8] で問題に割り当てられている番号である。

表 4.1 問題の大きさ

問題		大きさ	問題番号
実験 1 回目	A	10×10	8846
	B	7×7	7780
	C	10×10	1104
	D	15×15	10438
	E	10×10	5937
実験 2 回目	A	15×15	1146
	B	15×15	950
	C	15×15	799
	D	20×20	1059
	E	15×15	1094

4.3 比較評価の点数化

比較評価を問題の難易度の点数として実数値化した。まず点数化した比較評価 $\mathrm{compare}(X, Y)$ を式 4.1 で定義する。そして問題 X についての難易度の評価を評価点数 $\mathrm{score}(X)$ として式 4.2 で定義する。

$$\mathrm{compare}(X,Y) = \begin{cases} +1 & \text{if } X \neq Y \text{ and ``}X\text{ の方がとても難しい''} \\ +1 & \text{if } X \neq Y \text{ and ``}X\text{ の方が難しい''} \\ 0 & \text{if } X = Y \text{ or ``}X\text{ と }Y\text{ は同程度に難しい''} \\ -1 & \text{if } X \neq Y \text{ and ``}Y\text{ の方が難しい''} \\ -1 & \text{if } X \neq Y \text{ and ``}Y\text{ の方がとても難しい''} \end{cases} \tag{4.1}$$

$$\mathrm{score}(X) = \sum_{Y\in\{A,B,C,D,E\}} \mathrm{compare}(X,Y) \tag{4.2}$$

4.4 結果

データ収集実験の被験者数は 1 回目は 30 人（うち 2 人は問題の解答時間の未回答あり)、2 回目は 26 人だった。

データ収集実験で用いた問題から特徴量を抽出した。無作為更新の試行の回数は $N = 10^5$ とした。2 回目の実験に用いた問題から抽出された特徴量を表 4.2 に記す。

実験で用いた 5 問の難易度評価の点数と抽出した特徴量の相関係数を各被験者ごと計算した。表 4.3 および 4.4 に 2 回目の実験から一部抜粋して記す。最も左の列は被験者 1 人 1 人に割り当てた番号。太字はお絵かきロジックの経験が有ると答えた被験者。下線付きは強い相関（相関係数 +0.7 以上または −0.7 以下)。

5. 個人モデルの抽出

得られた相関係数について、平均終了ステップ数 $\overline{T^E}$ と正の強い相関がある被験者はおよそ半数であった。これは

表 4.2 抽出された特徴量 (実験 2)

特徴量	A	B	C	D	E
$\overline{d}(E)$	0.450	0.517	0.280	0.440	0.222
$\overline{d}(h_1)$	0.069	0.053	0.013	0.017	0.011
$\overline{d}(h_2)$	0.286	0.147	0.053	0.052	0.036
$\overline{d}(h_3)$	0.029	0.028	0.018	0.024	0.039
$\overline{d}(h_4)$	0.027	0.045	0.041	0.053	0.016
$\overline{d}(h_5)$	0.031	0.137	0.059	0.073	0.024
$\overline{d}(h_6)$	0.032	0.035	0.021	0.029	0.028
$\overline{T^E}$	59.39	50.72	76.07	77.53	83.73
$\sigma^2(T^E)$	27.02	26.92	23.02	46.44	23.80
$T^E_{\min}$	39	31	55	53	62
$T^E_{\max}$	85	76	95	111	104

表 4.3 データ収集実験の被験者の評価点数と特徴量の相関:解法出現率 (2 回目の実験の被験者の一部を抜粋、太字は経験者、下線付きは強い相関)

	$\overline{d}(E)$	$\overline{d}(h_1)$	$\overline{d}(h_2)$	$\overline{d}(h_3)$	$\overline{d}(h_4)$	$\overline{d}(h_5)$	$\overline{d}(h_6)$
1	-0.40	-0.21	-0.09	$\underline{0.77}$	-0.63	$\underline{-0.71}$	0.08
2	-0.15	0.14	0.13	$\underline{0.98}$	$\underline{-0.72}$	-0.34	0.52
3	$\underline{0.80}$	$\underline{0.91}$	$\underline{0.86}$	0.20	0.00	0.19	$\underline{0.87}$
4	$\underline{-0.87}$	$\underline{-0.92}$	$\underline{-0.89}$	0.22	-0.27	-0.31	-0.65
5	-0.45	$\underline{-0.84}$	$\underline{-0.78}$	0.03	0.16	-0.29	-0.47
6	-0.63	$\underline{-0.80}$	-0.69	0.21	-0.14	-0.57	-0.52
7	-0.36	0.06	0.17	$\underline{0.90}$	$\underline{-0.85}$	-0.69	0.25
8	0.15	-0.50	-0.59	-0.10	0.58	0.29	0.01
9	-0.69	$\underline{-0.72}$	-0.67	0.61	-0.44	-0.48	-0.25
10	-0.52	-0.06	0.09	$\underline{0.80}$	$\underline{-0.88}$	$\underline{-0.84}$	0.02
11	-0.66	-0.65	-0.52	0.49	-0.42	$\underline{-0.71}$	-0.35
12	$\underline{-0.81}$	$\underline{-0.70}$	-0.55	0.45	-0.51	$\underline{-0.77}$	-0.50

表 4.4 データ収集実験の被験者の評価点数と特徴量の相関:終了ステップ数 (2 回目の実験の被験者の一部を抜粋、太字は経験者、下線付きは強い相関)

	$\overline{T^E}$	$\sigma^2(T^E)$	$T^E_{\min}$	$T^E_{\max}$
1	0.55	0.16	0.54	0.58
2	0.07	-0.14	0.09	0.06
3	$\underline{-0.75}$	0.18	$\underline{-0.78}$	-0.53
4	$\underline{0.84}$	-0.16	$\underline{0.87}$	0.63
5	$\underline{0.86}$	0.59	$\underline{0.80}$	$\underline{0.97}$
6	$\underline{0.95}$	0.41	$\underline{0.91}$	$\underline{0.98}$
7	0.28	-0.17	0.31	0.23
8	0.36	$\underline{0.86}$	0.27	0.64
9	$\underline{0.82}$	0.10	$\underline{0.82}$	$\underline{0.76}$
10	0.44	-0.21	0.47	0.34
11	$\underline{0.88}$	0.25	$\underline{0.87}$	$\underline{0.88}$
12	$\underline{0.93}$	0.06	$\underline{0.93}$	$\underline{0.83}$

解き終わるまでの手数が多い問題ほど難しいと評価する人がおよそ半数であったことを意味する。これらの人達は $\overline{T^E}$ を基準とすることが難易度の推定において適している。その他の被験者は $\overline{T^E}$ と相関がないため、$\overline{T^E}$ に基いても良い難易度の推定はできない。

被験者 7 に注目する。表 4.4 から $\overline{T^E}$ と相関が弱く $\overline{T^E}$ に基づいた難易度の推定は適さない。しかし表 4.3 から $\overline{d}(h_3)$ や $\overline{d}(h_4)$ と相関が強いため、これらの解法出現率に基づいた推定が適していると考えられる。

被験者 8 に注目する。被験者 8 は本研究で設定した特徴量の内で唯一 $\sigma^2(T^E)$ にのみ強い相関があった。問題を一度解いただけでその問題を解くために必要な手数のばらつきを知ることはできない。あくまでも必要な手数にばらつきが出やすい問題について、難しいと感じる傾向があるといえる。

人によって強い相関のある特徴量が異なり、難易度の感じ方の原因が異なることが分かった。本研究で行った実験の全ての被験者について、難易度の感じ方に強い相関のある特徴量が見つけられた。個人に対応した難易度の推定のための基準とすることができる。

6. おわりに

お絵かきロジックの問題から取り出した使用可能な解法が出てくる割合と解く手数に注目した特徴量を取り出した。問題を実際に解いて 2 問ずつ一対で比較することで評価した難易度の感じ方を収集した。それらの特徴量と人それぞれの難易度の感じ方にある関係を相関として示すことができた。

回答者によって異なる難易度の感じ方は相関の強い特徴量の違いに現れた。強い相関を見せる特徴量を基準とすることで、回答者の難易度の感じ方を推測することができるようになる。これにより図 1.1 に表したような、個人に対応した難易度を推定するシステムを構築することができる。

同一の解法であってもマスの状態やヒントの数字の違いで、適用の難しさが異なることがある。これを考慮した特徴量の抽出も行いたい。また終了ステップ数を求める試行を回答者に依存しない方法で行った。実際には回答者が好んでよく使う解法で盤面を更新して終了ステップ数を求める必要がある。

参考文献

[1] 激辛数独 11. ニコリ, 2012.
[2] 伊藤. ヒューリスティックスを用いたロジックパズルの難易度自動評価. 2005.
[3] 小場, 中所. 数独の難易度判定アプリケーションの提案と評価. 2011.
[4] 乾, 小谷. ナンプレの解法, 難易度の算出, 問題の作成. 2002.
[5] いしだのん. ののぐらむ—絵が出てくる数理パズル. 日本評論社, 2005.
[6] イラストロジックベスト・オブ・ベスト 名作・傑作セレクト vol.1. 学研パブリッシング, 2015.
[7] ペイントロジック 2017 年 03 月号, 2017.
[8] お絵かきロジック, 2017. http://www.minicgi.net/logic/.

付　　録

A.1 発見的解法

h_1 **総和一致：** (行のマス数) = (ヒントの数字の総和) + (ヒントの数字の個数) - 1 が成り立つ時、端から可能な限り間を詰めて塗っていったただ 1 通りの塗り方にしか定まらず、その塗り方に確定する。

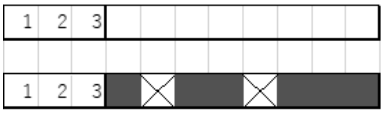


図 **A·1** 発見的解法：総和一致

h_2 **全ヒント確定：** 行のヒントの数字の使い方が全て定まっている場合、その行の全ての未確定マスは塗らないマスに確定する。

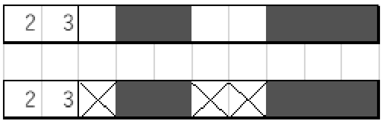


図 **A·2** 発見的解法：全ヒント確定

h_3 **両端の塗るマス注目：** 行の左（右）端から左（右）端のヒントの数字分のマスを 1 つの範囲として見て、その範囲の中に塗るマスがある場合、範囲の右（左）端から連続した未確定マスは塗るマスに確定する。

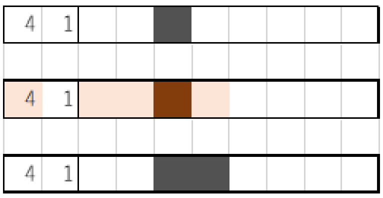


図 **A·3** 発見的解法：両端の塗るマス注目

h_4 **両端の塗らないマス注目：** 行の左（右）端から左（右）端のヒントの数字分のマスを 1 つの範囲として見て、その範囲の中に塗らないマスがある場合、範囲の左（右）端から連続した未確定マスは塗らないマスに確定する。

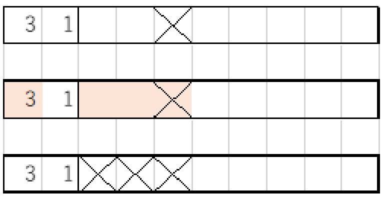


図 **A·4** 発見的解法：両端の塗らないマス注目

h_5 **最大ヒント注目：** 確定した塗りマスの連続数が行で最大のヒントに一致する場合、その連続の両隣 1 マスは塗らないマスに確定する。

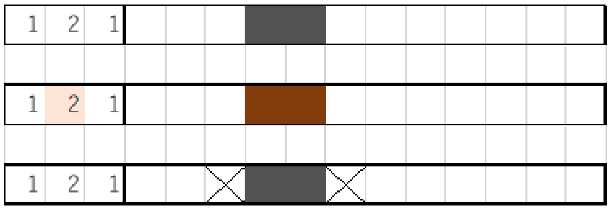


図 **A·5** 発見的解法：全ヒント確定

h_6 **最小ヒント注目：** 確定した塗らないマスに挟まれた未確定マスの連続数が行で最小のヒントよりも小さい場合、その連続の未確定マスは全て塗らないマスに確定する。

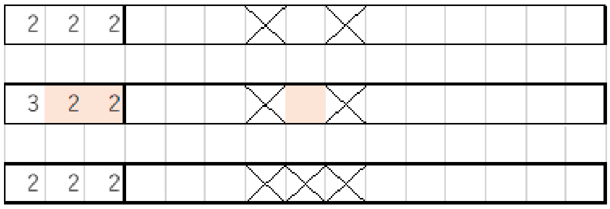


図 **A·6** 発見的解法：最小ヒント注目

トリミング： トリミングは他の解法と併せて使う。行の両端にある確定した塗りマスと塗らないマスについて、使い終わったヒントとそのヒントに対応した塗りマスを無視した行を新たに考える。こうして得た新しい行を切り落とされた行と呼ぶ。ある解法で切り落とされた行が更新可能であれば解法を適用して得た行を切り落とす前の行に反映する。

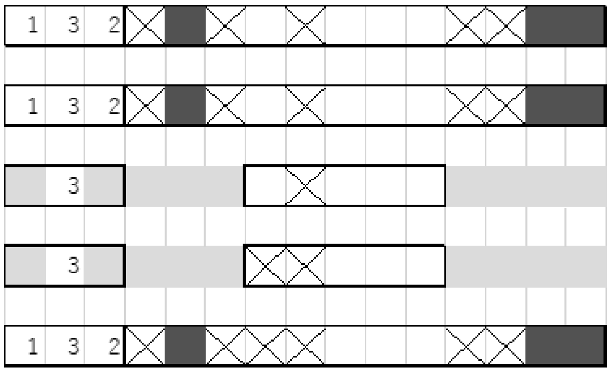


図 **A·7** トリミング付きの発見的解法