2018年　3月 7日

第13回組合わせゲーム･パズル研究集会

「四角に切れ」の定数時間検査アルゴリズム

電気通信大学
情報･ネットワーク工学専攻
伊藤･Belmonte研究室　竹田美弘

1. はじめに

一般的に、アルゴリズムが問題を解くとき
入力は全て読み込んで判断する必要がある。

巨大情報を扱う場合、一通り読み込むだけで
大変な時間がかかる…。

データ数に無関係な定数個のデータをみる

定数時間アルゴリズム

→計算時間は $O(1)$ になり、巨大情報も扱いやすい。

- 近年はビッグデータとの関係で注目を浴びている。

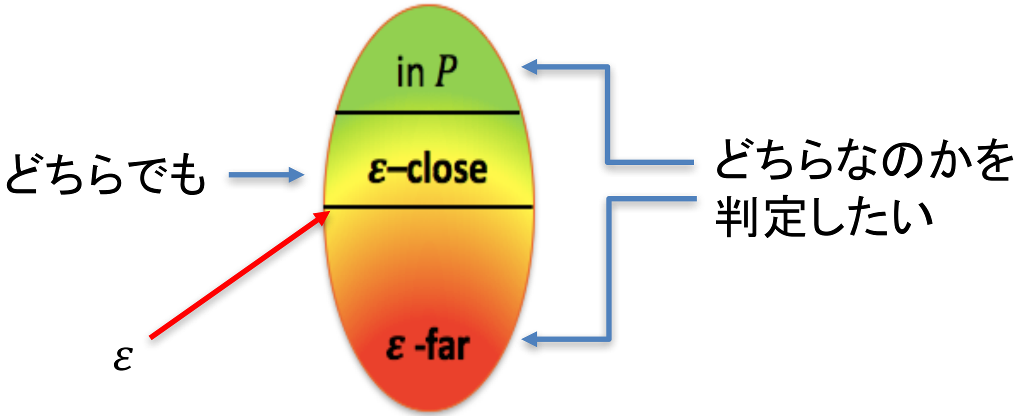
性質検査

定数時間アルゴリズムで最もよく使われる枠組み

⇒ **性質検査**

入力が所望の性質を
満たしているか否かを判定する
判定問題の緩和

入力と性質 Pとの間の
距離(0〜1に正規化する)
を定義し,その距離と
任意の実数 $\varepsilon\,(0 < \varepsilon \leq 1)$
とを比較して判定する



背景

- プログラムのチェック
- グラフ
- 関数の判定
- NP完全問題(ナップザック問題)
- ボードゲーム(将棋、チェス)

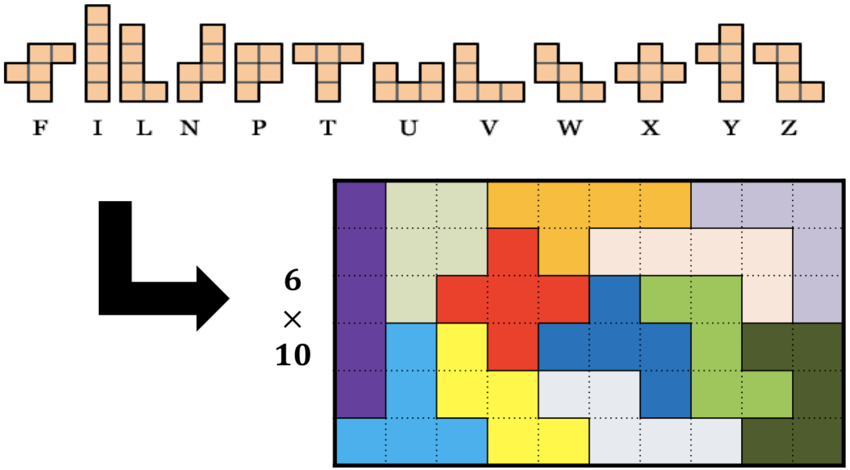
に対する
定数時間アルゴリズム

しかし、まだボードパズルに対するものは研究されていない...

⇒ **ボードパズル**に対して定数時間アルゴリズムの研究を行った

既存成果

- 前回の本研究集会にて紹介したポリオミノという詰め込みパズルは定数時間検査可能であることが示せた。
- この結果は国際会議であるJCDCG[3]にて発表も行った。

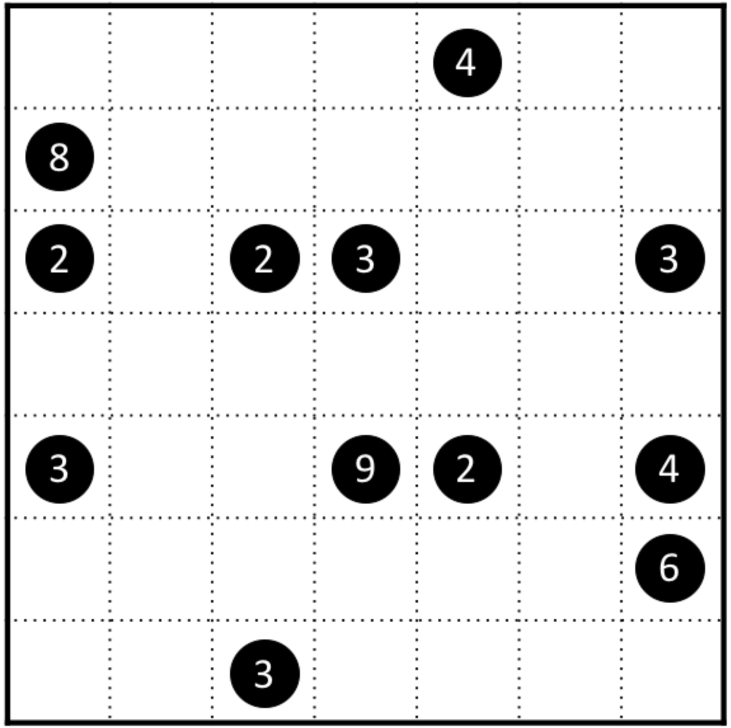


※ポリオミノと同様の手法を用いて
「四角に切れ」に対する定数時間アルゴリズムを
構築することができた。

2. 四角に切れの検査可能性

- **四角に切れとは**

盤面上にいくつかの
数字が記されたマスが存在する
任意のサイズの正方形の
盤面が与えられる。

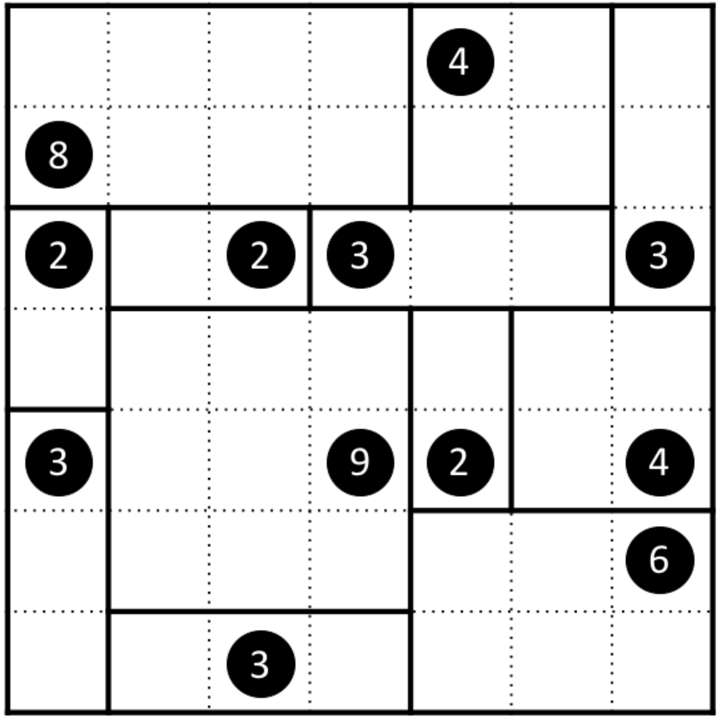


2. 四角に切れの検査可能性

- **四角に切れとは**

盤面上にいくつかの
数字が記されたマスが存在する
任意のサイズの正方形の
盤面が与えられる。

これを数字をちょうど一つずつ含み、
その値と等しい面積の長方形
（正方形も含む）に分割することが
できるか否かを判定するパズル。



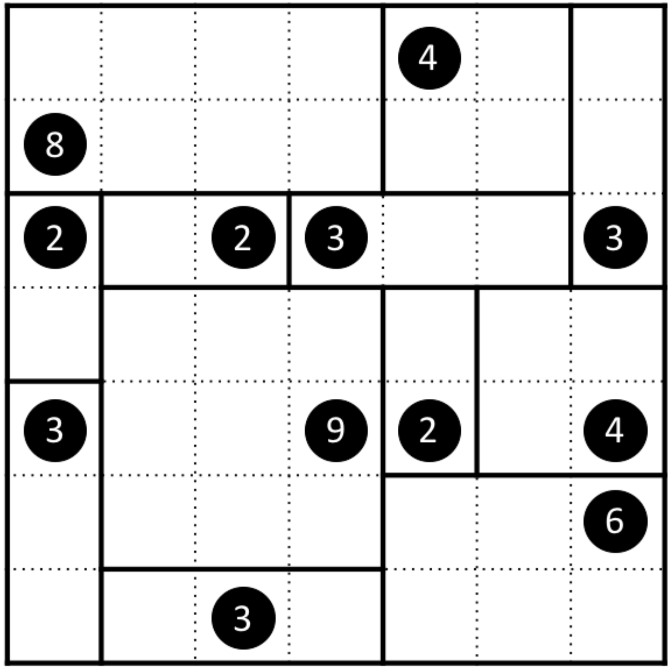
四角に切れとは - 用語

本研究において、便宜上、

数字の記されたマスを数字マスと呼び、
各数字マスに記された値は
定数である k 以下であるとする。

また、

あるマスが分割の際長方形に含まれるとき、
そのマスが覆われると呼ぶこととする。



諸定義 - 四角に切れ

四角に切れ

- 問題例(入力)

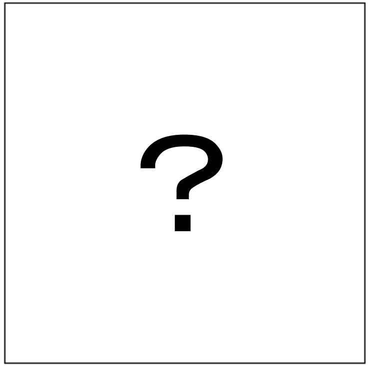
面積 n の盤面 S と
数字マスの座標の組 $i, j \in \{1, 2, \dots, \sqrt{n}\}$ とその値

- 要請

S のすべてのマスを、各数字マスをちょうど1つずつ含む、その値と等しい面積の長方形で分割することが可能か否かを判定せよ。

※この要請の答えが「可能」であるならばその問題例は**実行可能**であるという

諸定義 - 座標オラクル



- **入力**

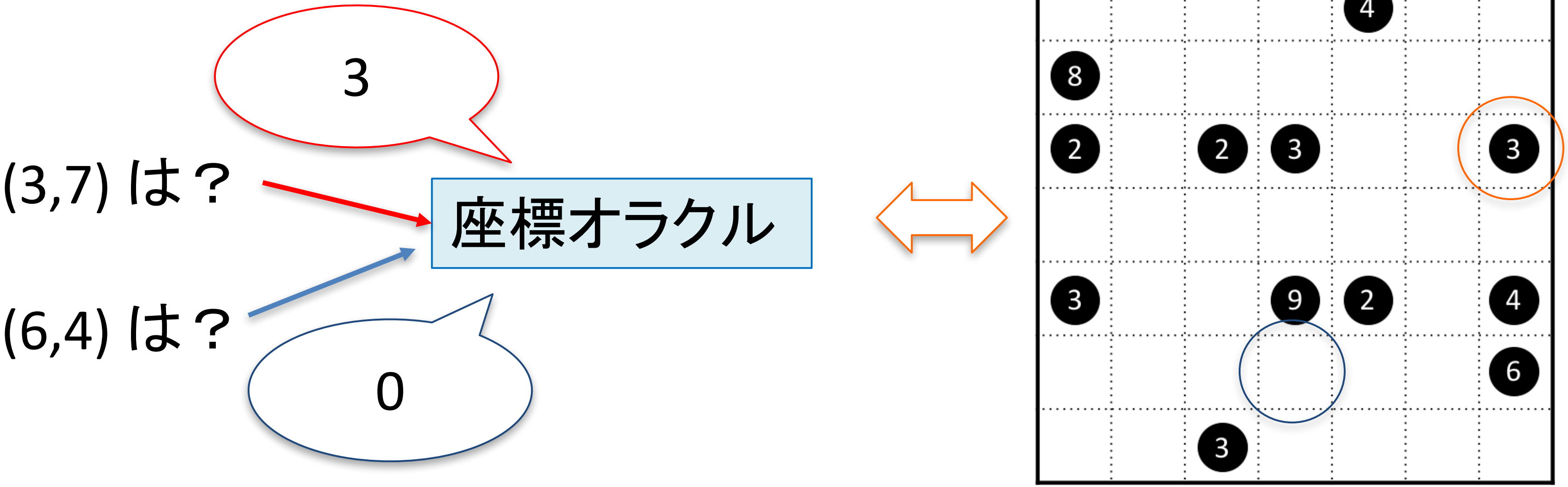
  盤面上にある
  各数字マスの座標およびその値はいくらか？

⇨ これを得るために、アルゴリズムは以下のオラクルを持つ

- 座標オラクル

  任意の整数の組 $i, j \in \{1, 2, \ldots, \sqrt{n}\}$ を与えることにより、
  座標 (i, j) にあるマスの数値（記されていない場合は0）を
  $O(1)$時間で回答する。

諸定義 - 質問計算量



※ 上述したオラクルに対し、問題例Qを明示するときには $q(i,j;Q)$ と表す

アルゴリズムは問題を解くために、座標オラクルに何度か質問を行う。その回数を**質問計算量**という

諸定義 - 問題例の距離

盤面上のマスに
どの程度の違いがあるか

このとき、二つの問題例 Q と Q′ の**距離** dist(Q, Q′) を
答えの異なる座標オラクルの総数をnで除したものとする。

$$\mathrm{dist}(\mathrm{Q},\mathrm{Q}') = \frac{|\{i,j \mid q(i,j;Q) \neq q(i,j;Q')\}|}{n}$$

	4			
			6	
	4			
	2			
3		2		4

全25マス中赤い枠の
2マスのみが異なる

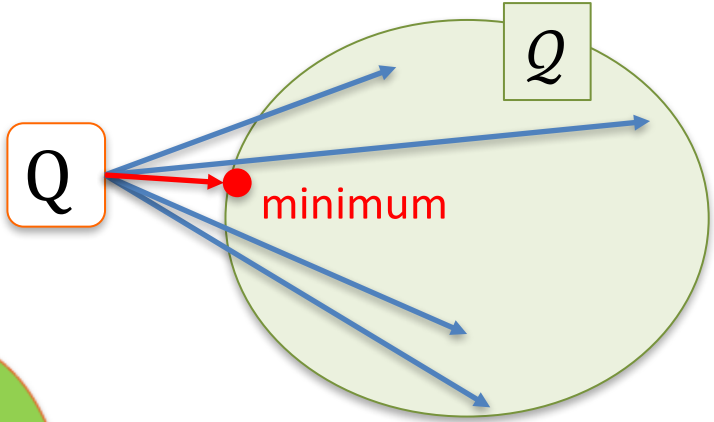
距離は 2/25

	4			3
			3	
	4			
	2			
3		2		4

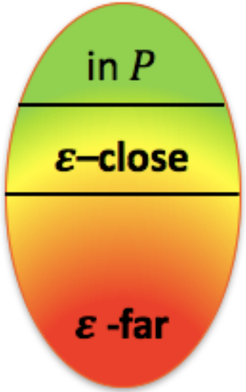
諸定義 - 問題例の集合との距離

さらに、任意の問題例 Q と任意の問題例の集合 $\mathcal{Q}$の距離は次のように置く。

$$\mathrm{dist}(\mathrm{Q}, \mathcal{Q}) = \min_{Q' \in \mathcal{Q}} \mathrm{dist}(\mathrm{Q}, \mathrm{Q}')$$

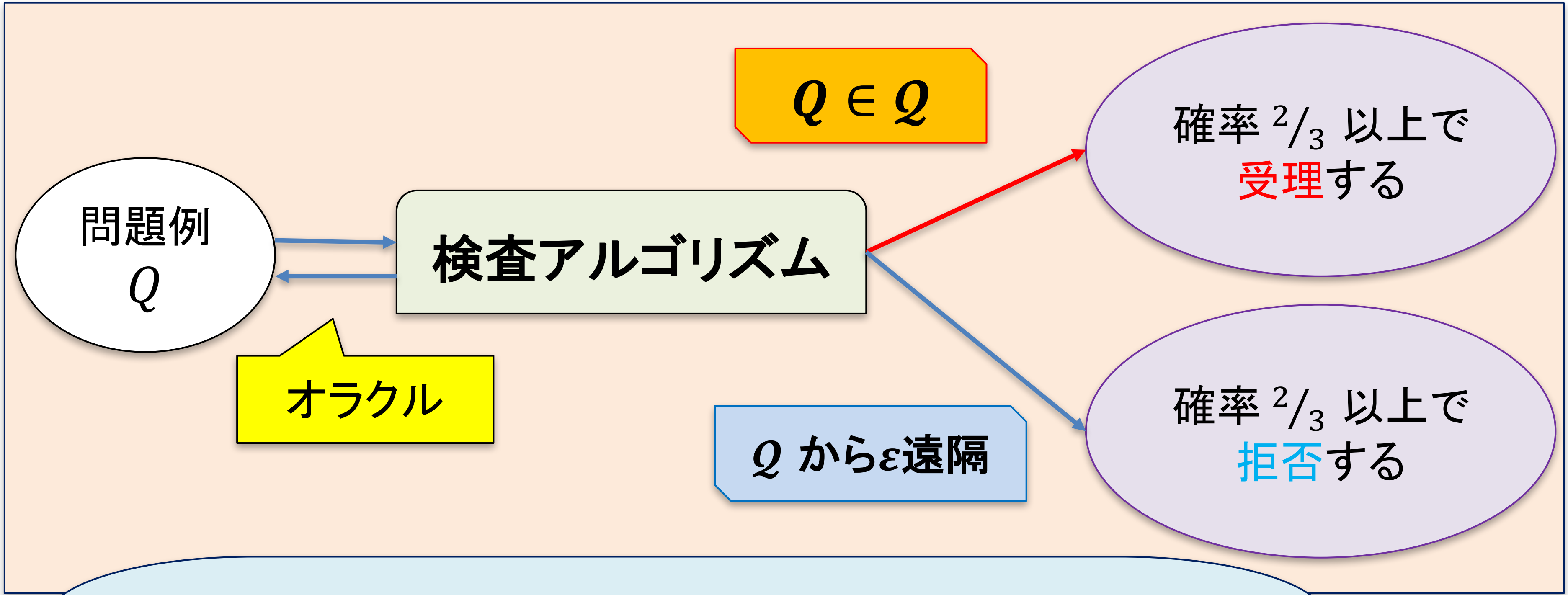


任意の $0 < \varepsilon \leq 1$ に対して、
$\mathrm{dist}(Q, \mathcal{Q}) \geq \varepsilon$ のとき
ε遠隔(ε-far)であるといい、
$\mathrm{dist}(Q, \mathcal{Q}) < \varepsilon$ のとき
ε近接(ε-close)であるという。



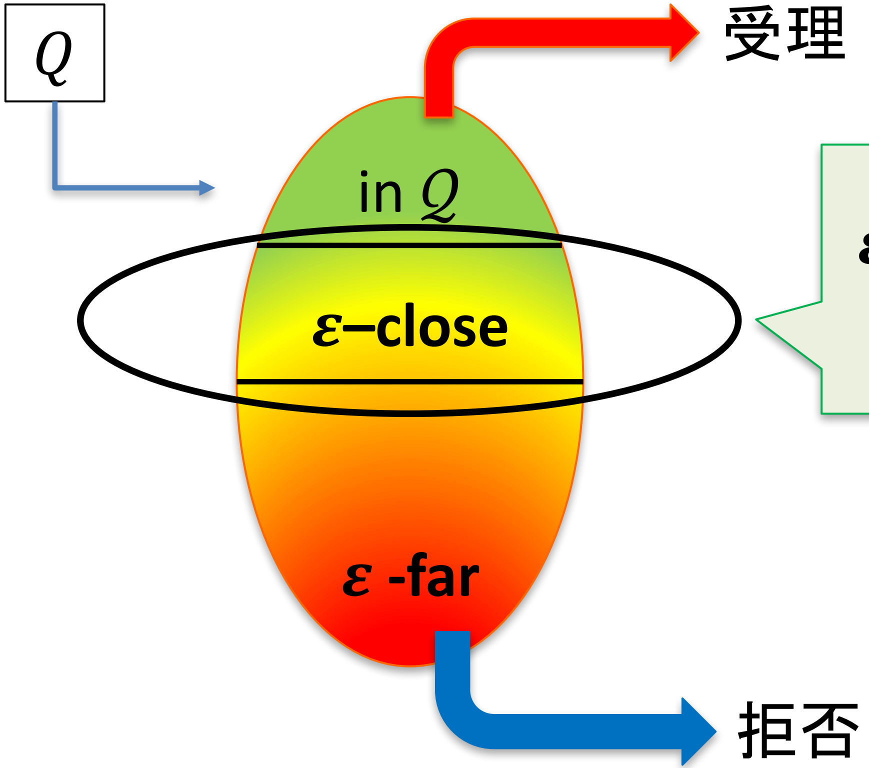
諸定義 - 検査アルゴリズム

$\mathcal{Q}$ に対する検査アルゴリズム



※質問計算量が定数ならば**検査可能**であるという

検査アルゴリズムの補足

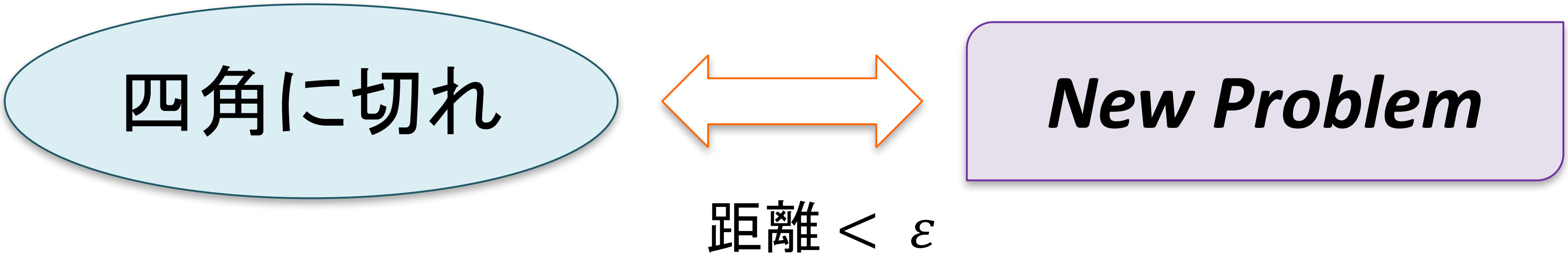


ε未満のわずかな誤差なので
許容して受理しても良い

四角に切れの解法とその検査

問題:四角に切れ を定数時間で解くために

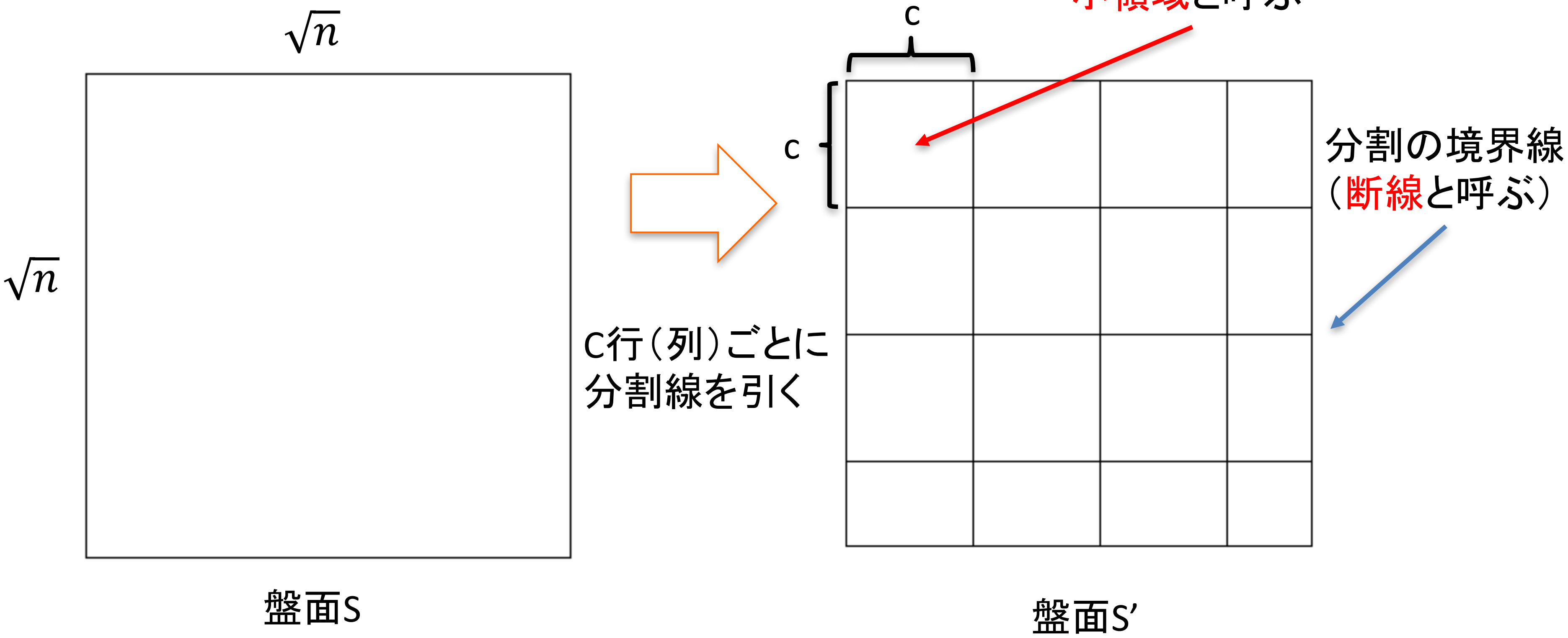
→　距離の近い、違う形の問題を新たに定義する。



盤面の分割

盤面を$c \times c$ (c: εに依る定数) に分割することを考える

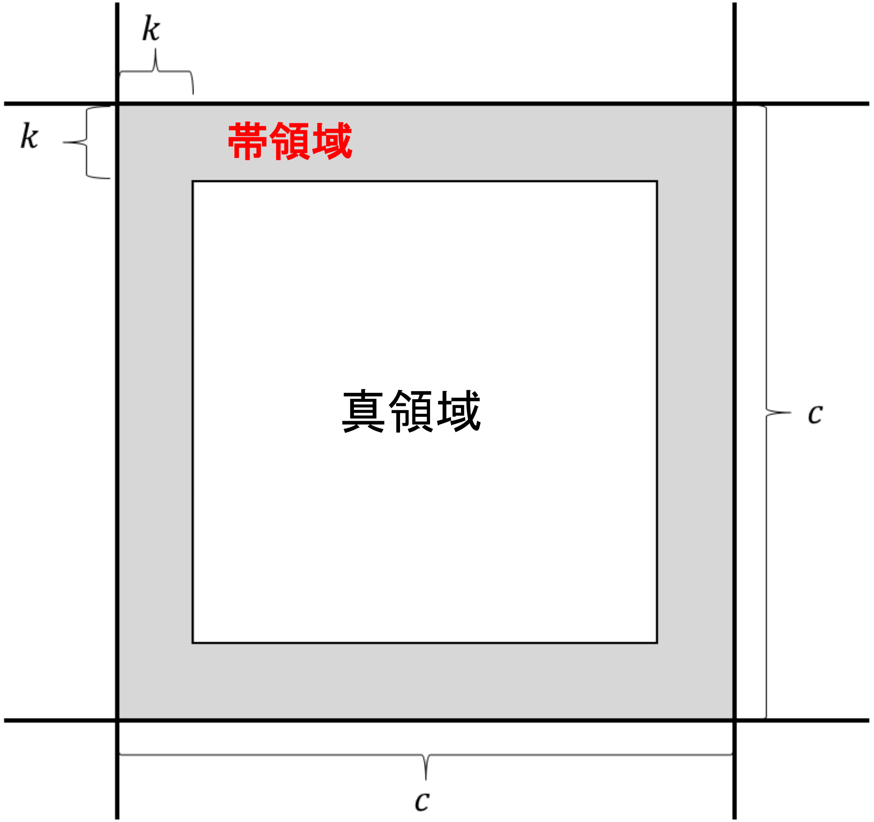
この $c \times c = c^2$
の分割盤面を
小領域と呼ぶ



諸定義 - 小領域の分類

小領域内に新たな領域を定義する。

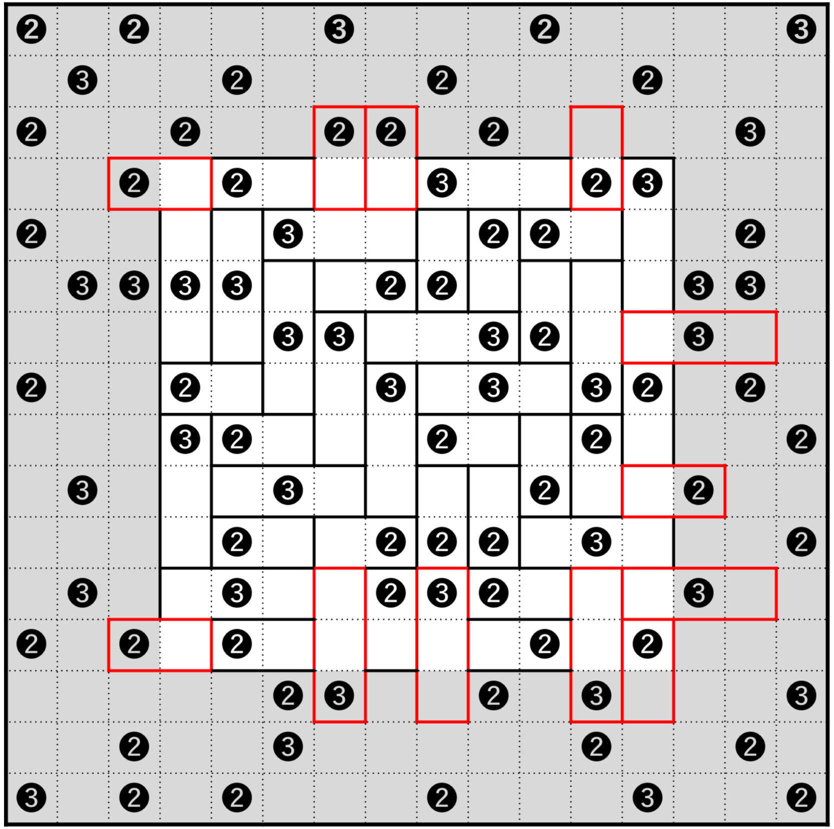
- 図が一つの小領域を表す。
- 断線から k 分の範囲を帯領域と呼ぶ。（図の灰色部）
- それ以外を真領域と呼ぶ。（図の白色部）



諸定義 - 擬カッティング

帯領域の数字マスに関しては
使っても使わなくてもよいとした上で、
本来と同様に真領域内のマスが
全て覆われるように
長方形で分割することを
擬カッティングと呼ぶ。

※ 例えば右図の赤色部
のような長方形の
取り方をしても良い



分割問題の定義

分割四角に切れ

- 問題例(入力)

面積 n の盤面 S と
数字マスの座標の組 $i, j \in \{1, 2, \ldots, \sqrt{n}\}$とその値

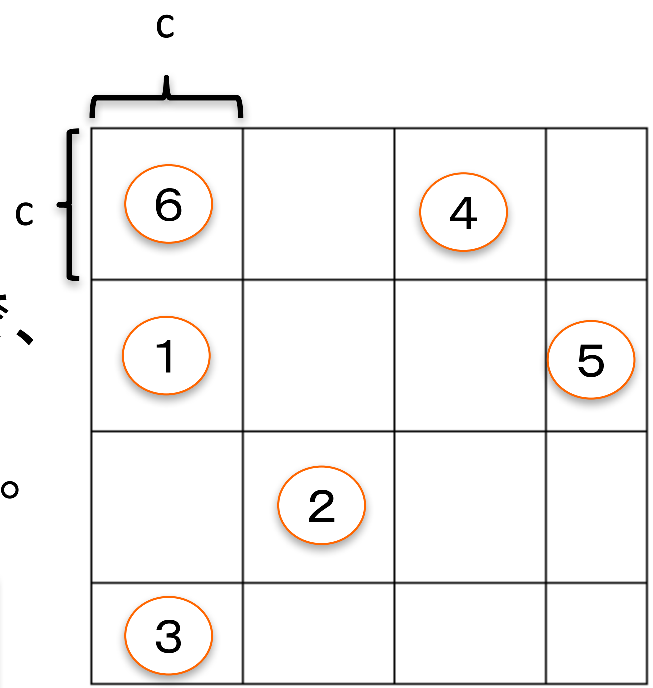
- 要請

盤面を $c \times c$の小盤面に区切り、小盤面を分類した上で、
S のすべての小領域を擬カッティングすることが
可能か否かを判定せよ。

分割四角に切れの解法

◎ 小領域を一様ランダムにm個選び、
それぞれ擬カッティング可能か否か
判定する。

- 小領域の面積は高々 $c \times c = c^2$ なので、合計でも mc^2 回オラクルに質問すればすべての情報を得た上で判定できる。



$\mathrm{m} = \dfrac{\log 6}{2\varepsilon'^2}$ とおけば、高い確率で
全ての小領域が擬カッティング可能で
あるかを誤差 ε' 未満で推定できる

分割四角に切れの検査アルゴリズム

- 次のアルゴリズムによって、分割四角に切れは検査可能である

Procedure TESTINGDIVIDED-SHIKAKU

begin

1. $n \leq \frac{\mathrm{c}^2 \log 6}{2\varepsilon'^2}$ の場合、オラクルで全ての情報を得た上で判定し、終了する。
2. $n > \frac{\mathrm{c}^2 \log 6}{2\varepsilon'^2}$ の場合、一様ランダムに $\mathrm{m} = \frac{\log 6}{2\varepsilon'^2}$ 個の小領域を選び、それぞれ小領域内の全ての座標に座標オラクルで質問した上で擬カッティング可能か否か判定する。全ての小領域が擬カッティング可能であれば受理し、そうでなければ拒否する。

end.

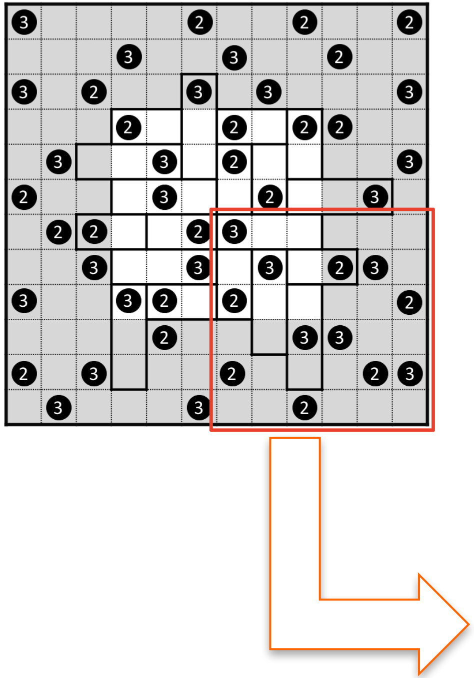
四角に切れの検査可能性

- 従って、次の補題を証明することができた。

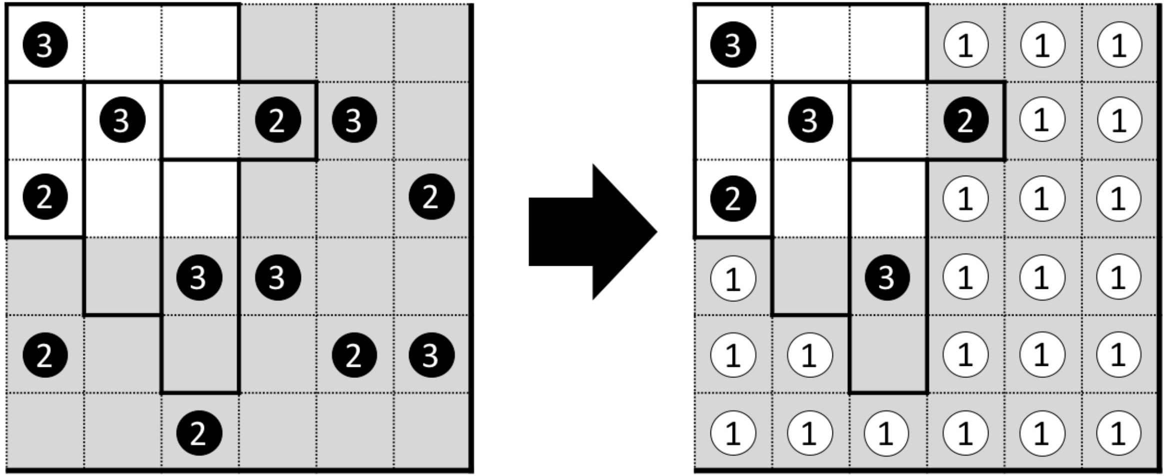
補題

任意の正整数 $c \geq 1$, 任意の正実数 $\varepsilon' > 0$ に対し、
分割四角に切れには質問計算量が
$O(c^2 \varepsilon'^{-2})$ である検査アルゴリズムが存在する。

四角に切れとの距離 - 擬カッティング後の置き換え



- 擬カッティングの際に、覆われなかったマスが帯領域にほとんどの場合存在する。
- このようなマスについては、後から全て数値1を持つ数字マスに置き換える。

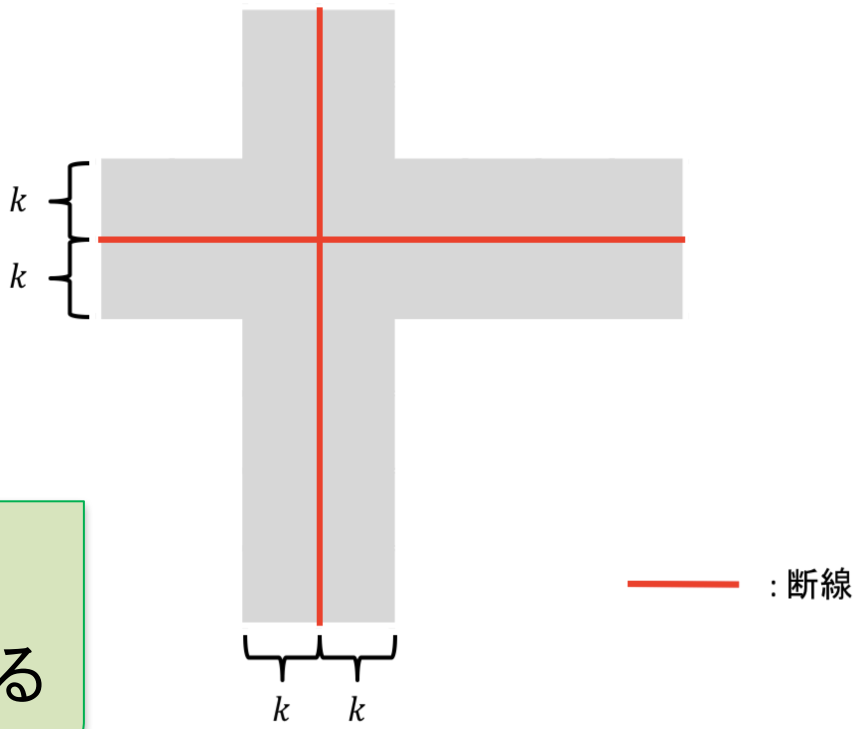


四角に切れとその分割問題との距離

- これによって変化の生じるマスは各小領域における帯領域にあたる部分のみである。

- 該当するマスは盤面全体でも高々 ${}^{4kn}/_{c}$ 個

$c < {}^{4k}/_{\varepsilon''}$とおけば、
ここで生じる距離は ε''で抑えられる



四角に切れの検査アルゴリズム

- 次のアルゴリズムによって、四角に切れは検査可能である

Procedure TESTINGSHIKAKU

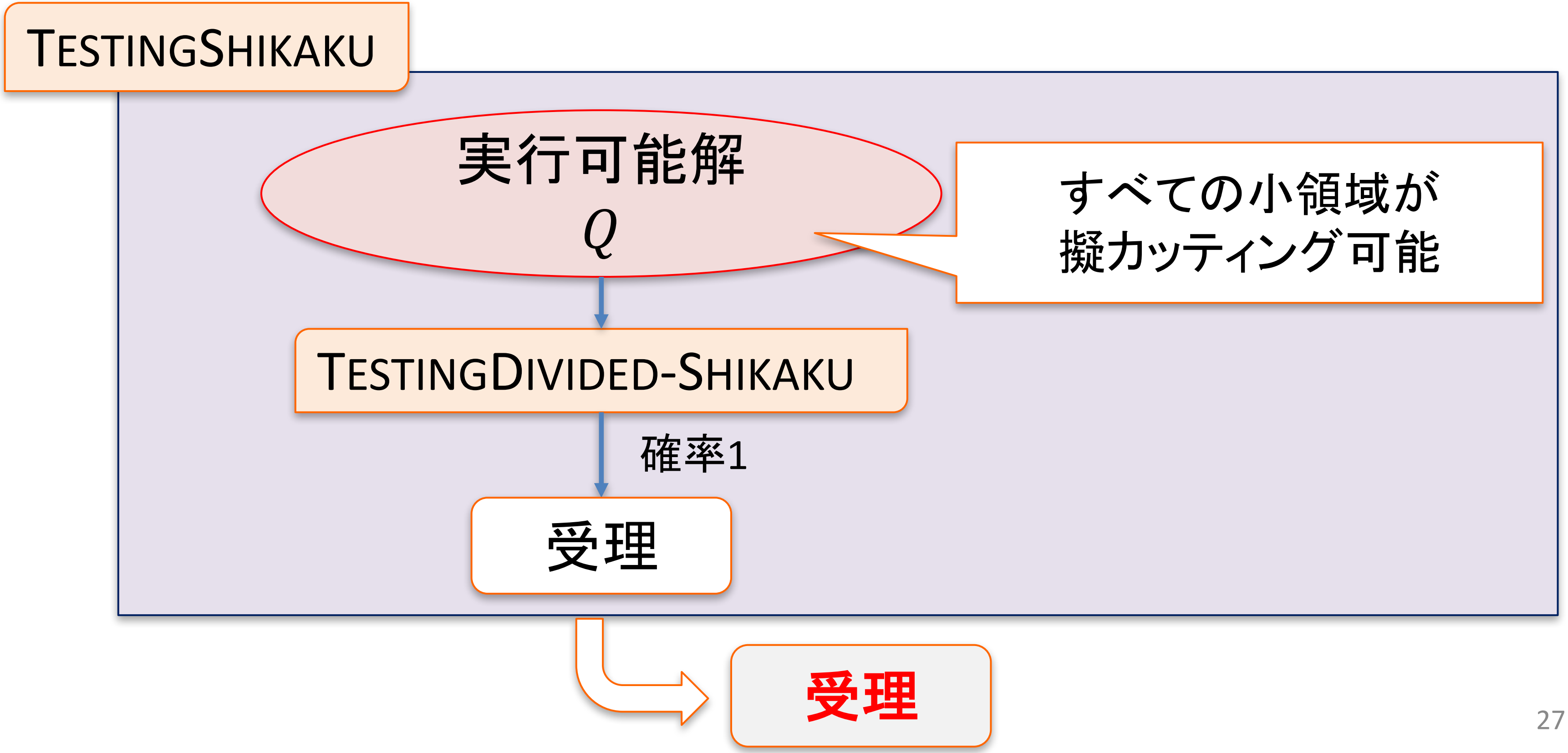
begin

1. $n \leq \frac{2c^2 \log 6}{\varepsilon^2}$ の場合、オラクルで全ての情報を得た上で判定し、終了する。
2. $n > \frac{2c^2 \log 6}{\varepsilon^2}$ の場合、$\varepsilon' = \varepsilon'' = \frac{\varepsilon}{2}, c = \frac{8k}{\varepsilon} + 1$ としたうえで、TESTINGDIVIDED-SHIKAKUによって Q が分割四角に切れにおいて実行可能であるか、それから ε'遠隔であるかの検査を行い、Qが受理されたら受理し、拒否されたら拒否する。

end.

四角に切れの検査アルゴリズムの正当性

- 正しく受理されることを示す

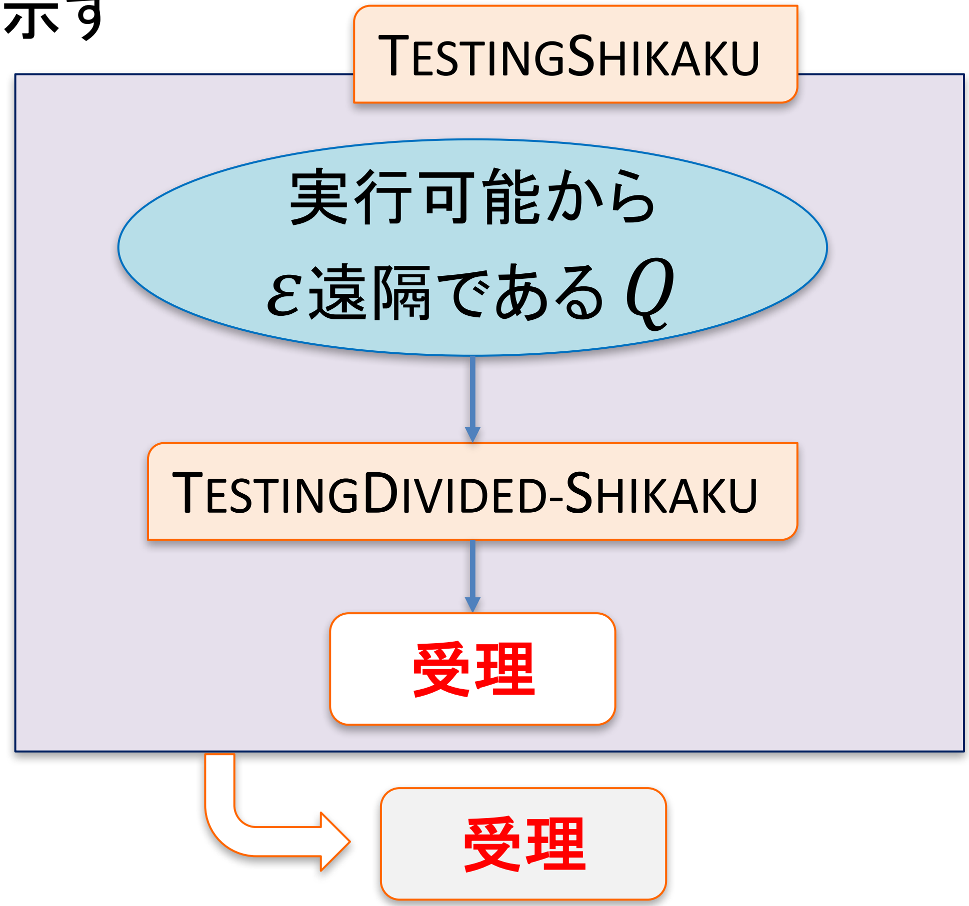


四角に切れの検査アルゴリズムの正当性

- 正しく拒否されることを示す

仮定

Qが確率2/3以上で
拒否されないとする



検査可能性

すなわち

Q はある分割四角に切れの実行可能解Q'に ε' 近接である

$\mathrm{dist}(Q, Q') < \varepsilon'$

$\mathrm{dist}(Q, Q'') < \varepsilon' + \varepsilon'' = \varepsilon$

実行可能に ε 近接

矛盾

さらに、

分割四角に切れの実行可能解Q'と四角に切れの実行可能解Q''の距離は高々 ε''で抑えられる

$\mathrm{dist}(Q', Q'') < \varepsilon''$

四角に切れの検査可能性

- 従って、次の定理を証明することができた。

定理

四角に切れには質問計算量が
$O(k^2 \varepsilon^{-4})$ である検査アルゴリズムが存在する。
すなわち、四角に切れは検査可能である。

4. 結果

◎ まとめ

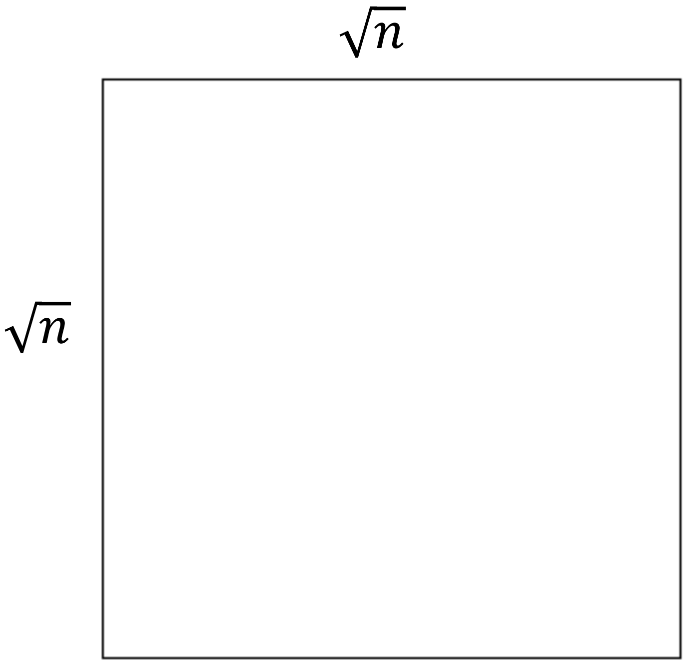
- ライツアウト、ポリオミノ、四角に切れに対して、定数時間検査可能なアルゴリズムを与えることができた。
- 今回用いた盤面を定数サイズに分割する技法を用いれば、さらに多くのボードパズルについて検査可能となることが期待できる。

◎ 今後の課題

- 他のボードパズルの定数時間検査可能性を証明する。
- 検査不可能なボードパズルが存在するのか。
- 検査可能であるものとそうでないものとの汎用性の高い特徴づけを導く。

盤面の分割

盤面を定数サイズに分割することを考える

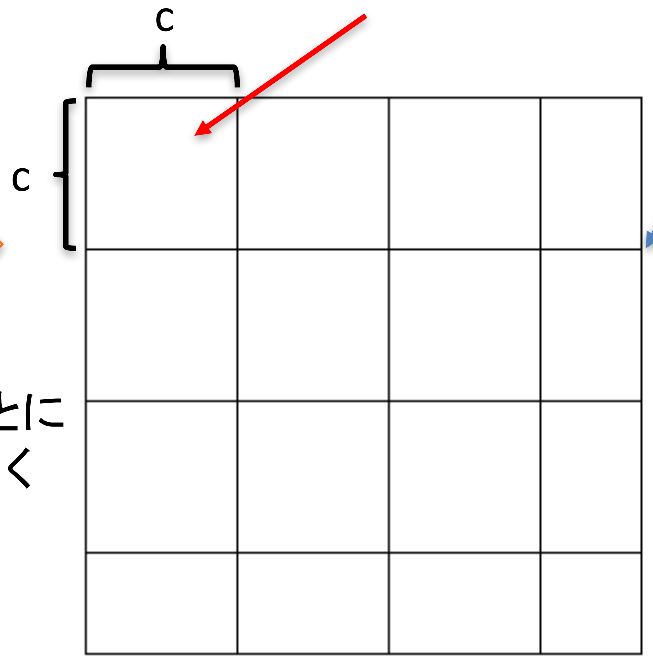


盤面S

C行（列）ごとに
分割線を引く

この $c{\times}c = c^2$
の分割盤面を
小領域と呼ぶ

分割の境界線
（断線と呼ぶ）

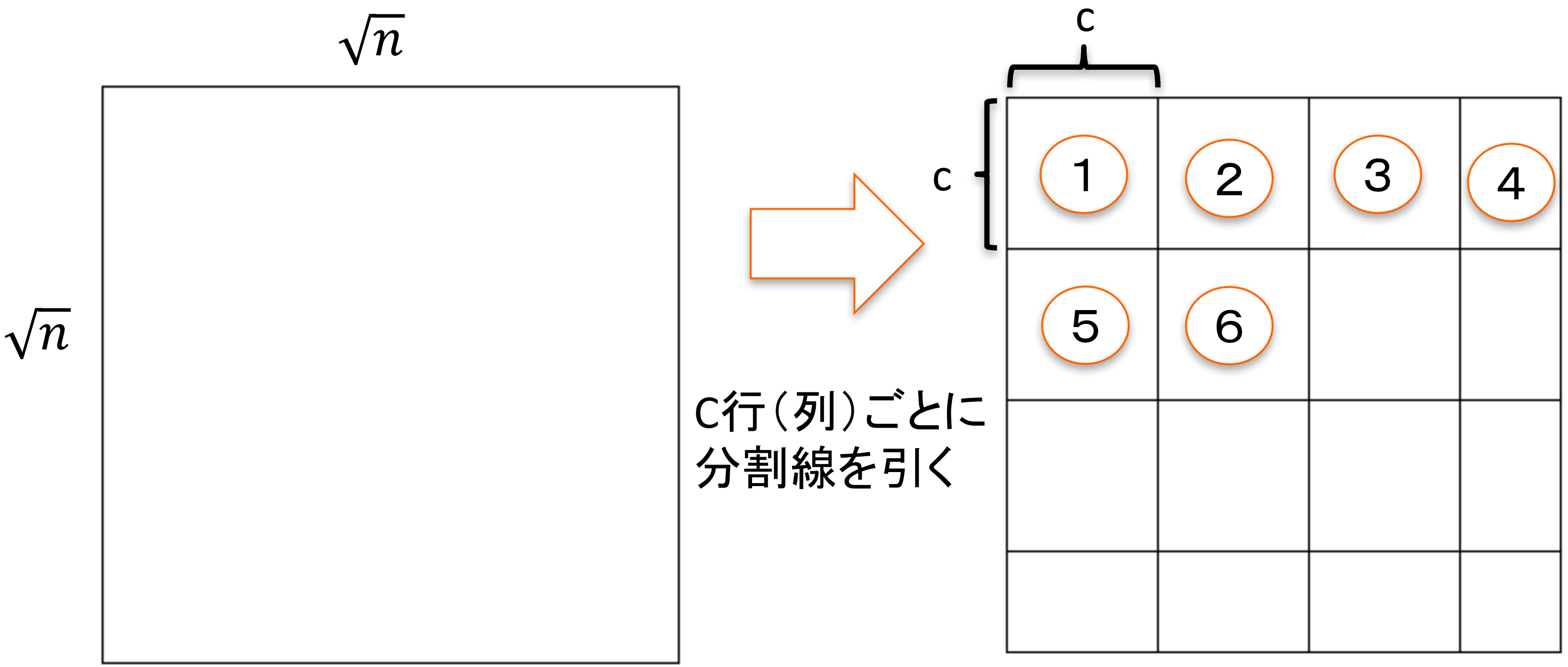


盤面S’

盤面の分割

盤面を定数サイズに分割することを考える

- 小領域に番号を振る（区別する）



盤面S

盤面S'

サンプリングのための諸定義

ランダムに小領域をm個選び、
ある小領域 iが擬カッティング可能ならば1を
そうでなければ0をとるような変数を$\mathrm{X^i}$ $(i = 1, \ldots, m)$とする。
また、$\overline{\mathrm{X}} = \frac{1}{m}\left(\mathrm{X^1} + \mathrm{X^2} + \cdots + \mathrm{X^m}\right)$とする。

- 小領域の面積は $c \times c = c^2$ なので
  合計でも mc^2 回の質問で済む
- 小領域内の情報をすべて得た上
  で小領域で解けるか調べる

⇒ もし解けないなら全体でも解けないので拒否する

サンプリングの保証

このとき Hoeffding の不等式より 任意の実数$\varepsilon' > 0$に対して

$$\Pr\left[\left|\overline{X^i} - E\left[\overline{X^i}\right]\right| \geq \varepsilon'\right] \leq 2\exp(-\frac{2m^2\varepsilon'^2}{m})$$

これが $^1/_3$ より小さくなれば良いから

$2\exp(-2m\varepsilon'^2) < {}^1/_3$ とすればよく、

$$m > \frac{\log 6}{2\varepsilon'^2}$$

とすれば誤差がε' 以上になる確率が$^1/_3$未満となる。

解法 - 擬カッティング後の置き換え

これによって変化の生じるマスの最大数は、
各小領域の帯領域にあたる範囲のみである。
これは各断線に対し、
その左右(上下)k マス分、合計2k マスである。
断線の長さは$\sqrt{n}$であるので、
断線一つで高々2k $\sqrt{n}$個の
ピースオラクルが変化する。
断線の数は全部で高々$\frac{2\sqrt{n}}{c}$であるので、
全体で変化するマスの数は高々 $2\mathrm{k}\sqrt{n}\times\frac{2\sqrt{n}}{c}=\frac{4\mathrm{kn}}{c}$となる。
従って、$\frac{4\mathrm{k}}{\mathrm{c}}<\varepsilon'''$とすればよい