

信州大学工学部

学士論文

誤り訂正符号としてのラテン方陣の復号アルゴリズム  
に関する基礎的考察

指導教員 西新 幹彦 准教授

学科 電子情報システム工学科  
学籍番号 18T2115G  
氏名 長嶋 凌

2022年2月28日

# 目次

1	はじめに	1
2	$n \times n$ ラテン方陣と問題設定	2
3	ラテン方陣の復号アルゴリズム	4
4	性能の検証	6
4.1	検証方法	6
4.2	結果と考察	9
5	まとめ	12
	謝辞	12
	参考文献	12
	付録 A ソースコード	13
A.1	ラテン方陣の復元成功, 失敗を数えるプログラム	13

# 1 はじめに

現代社会において、インターネット通信、電波通信における情報のやり取りは非常に大切である。シャノンは現実の通信の流れを図1のようにモデル化した[1]。送信者の送ったメッセージは、符号器によって符号化され通信路を通る。通信路から出てきた受信語は、復号器によって復号され受信者に伝わる。通信路で雑音の影響を受ける場合、送信された符号語と受信語は同じものとは限らないが、受信者には正しいメッセージが伝わるようにしたい。この過程全体は、情報源符号化と通信路符号化に分けられる。本研究では通信路符号化について考える。通信路符号化の目的は、受け取った受信語から送信された符号語を推定することである。このことを誤り訂正という。誤り訂正の方法は一通りではなく、さまざまな工夫の仕方が考えられる。

過去の研究では、数独というものに着目した研究がある[3]。数独はパズルの一種であり、部分的に数字が入っている表をヒントに、空白の部分を埋めることができが目的である。この手順は受信語から、送られた符号語が何であったかを当てる手順と同じである。このように数独は誤り訂正符号の一種であると考えることができる。また、数独と似ているものとしてラテン方陣というものがある。ラテン方陣は、一定のルールで数字が埋まった表のことであり、もしいくつかの数字が消えてもルールに従って復元できる場合がある。よって、ラテン方陣も数独と同様に誤り訂正符号と考えることができる。ラテン方陣を誤り訂正符号として用いた従来研究[4]がある。そこで提案された復号方法の性能は、既存の復号方法を超えていない。

本研究では、ラテン方陣の復号方法を考える中で、従来研究よりも性能の高い復号方法を見た。その基本的な考え方は、ルールに違反している数字をルールに従ったものに書き換えることである。シンプルな考え方であることから、他の符号にも応用できると考えられる。

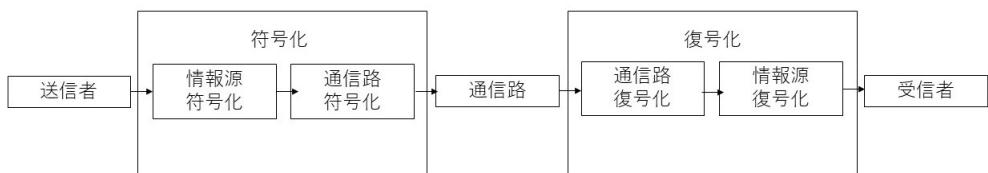


図1: シャノンによる通信のモデル

## 2 $n \times n$ ラテン方陣と問題設定

$n \times n$  ラテン方陣とは、 $n$  個の異なる数字が各行及び各列に 1 回だけ現れるよう各セルに配置された  $n \times n$  格子のことである。 $4 \times 4$  ラテン方陣の例を図 2 に示す。

1	2	3	4
2	1	4	3
3	4	2	1
4	3	1	2

図 2:  $4 \times 4$  ラテン方陣

本研究では、 $4 \times 4$  のラテン方陣を符号語として誤り訂正をしたい。それを図 3 に示している。

ラテン方陣の符号化レートを考える。一般に符号化レート  $R$  は、

$$R \triangleq \frac{\log_2 |M_n|}{\log_2 |\mathcal{X}^n|}$$

で定義される。ここで、 $M_n$  は符号、 $|M_n|$  は符号語数であり、 $|M_n| = 576$  である。 $\mathcal{X}$  は通信路アルファベットであり、本研究の場合、 $\mathcal{X} = \{1, 2, 3, 4\}$  である。 $|\mathcal{X}^n|$  は、アルファベットサイズであるためこの通信路では  $|\mathcal{X}^n| = 4^{16}$  となる。また、符号語長  $n$  は  $n = 16$  である。よって、符号化レート  $R$  は  $R = 0.287$  となる。

本研究では、 $4 \times 4$  ラテン方陣に合わせて 4 元対称通信路を想定する。具体的には、入力シンボルに対する出力シンボルの分布が図 4 のように表される通信路を考える。パラメータ  $\epsilon$  に対して、送信された符号語が雑音のある通信路を通り、正しくそのまま出力される確率を  $1 - \epsilon$  とし、誤った数字が出力される確率をそれぞれ等確率にしたいために、 $\frac{\epsilon}{3}$  とする。パラメータ  $\epsilon$  を雑音の強さと呼ぶことにする。

この通信路の通信路容量を確認する。通信路に対して、達成可能なレートの上限が通信路容量になるため、通信路容量は

$$C(W) \triangleq \max_X I(X; Y)$$

と定義される。 $I(X; Y)$  は  $X$  と  $Y$  の相互情報量である。相互情報量の定義より、

$$C(W) = H(Y) - H(Y|X)$$

のようにエントロピーを用いて表現することができる。そして、これを計算すると、

$$C(W) = 1 - \{(1 - \epsilon) \log_4 \frac{1}{1 - \epsilon} + 3\epsilon \log_4 \frac{1}{\epsilon}\}$$

となる。通信路符号化定理より、 $C(W) = 0.287$  を  $\epsilon$  について解くことによって、符号化レートが  $R = 0.287$  であるような符号の中で、雑音の強さが  $\epsilon = 0.149$  の通信路に対して復号誤り確率を限りなく小さくできる符号が存在することがわかる。つまり、ラテン方陣と同じ符号化レートをもつ理想的な符号は、 $\epsilon < 0.149$  の通信路に対して復号誤り確率がゼロになる。

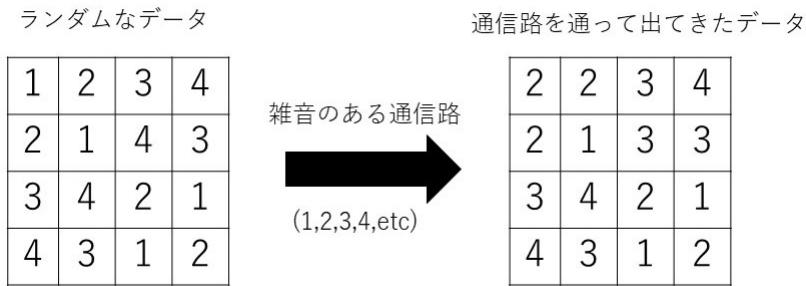


図 3: 通信路のモデル

入力 \ 出力	1	2	3	4
1	$1 - \epsilon$	$\epsilon / 3$	$\epsilon / 3$	$\epsilon / 3$
2	$\epsilon / 3$	$1 - \epsilon$	$\epsilon / 3$	$\epsilon / 3$
3	$\epsilon / 3$	$\epsilon / 3$	$1 - \epsilon$	$\epsilon / 3$
4	$\epsilon / 3$	$\epsilon / 3$	$\epsilon / 3$	$1 - \epsilon$

図 4: 条件付確率の表

### 3 ラテン方陣の復号アルゴリズム

ここでは、本研究で考えたラテン方陣の復号アルゴリズムについて説明する。通信路を通して出てきた値に対して、まずその値全体を確認する。そして行と列で同じ値が重複していた時は、ラテン方陣の条件を満たさない。そこで、ラテン方陣の条件を満たすようにするために重複している値については、1～4のどの数字が入るかわからないので、1～4のどの値も取れると考えることにする。その状態を図5に示す。

ここで、ラテン方陣の条件を満たす値を決定していくために図6に示すように1～4の値を二進数に置き換え、決まっていない数値は1111(15)とする。そして、ラテン方陣の決まっていない数値の同じ行と列にある値を確認する。図7のように決まっていない数値(1111)と他の同じ行や列にある数値(二進数)を反転したものとの論理積をとることで、決めたい数値の選択肢を消去していくことができるようなアルゴリズムである。ただし、雑音の強さが大きくなることで空欄が多くなる場合、復号アルゴリズムを通しててもラテン方陣が完成しないことがある。その場合は、復号失敗と同様の扱いとする。

1	2	1	4
2	1	4	3
3	4	2	1
4	3	1	2

	2		4
2	1	4	3
3	4	2	1
4	3		2

12 34	*2 **	12 34	* *4
*2 **	1* **	** *4	* 3*
1* 3*	** *4	*2 **	1* **
** *4	1* **	12 34	*2 **

図5: ラテン方陣の変換

1	1	1	1	15
0	0	0	1	1
0	0	1	0	2
0	1	0	0	4
1	0	0	0	8

$$1 << (n-1)$$

$$2^{(n-1)}$$

図6: 二進数変換

$$\begin{array}{r}
 1\ 1\ 1\ 1\ (15) \\
 \& 1\ 1\ 1\ 0\ (\sim 1) \\
 \hline
 1\ 1\ 1\ 0\ (14)
 \end{array}
 \quad
 \xrightarrow{\hspace{1cm}}
 \quad
 \begin{array}{r}
 1\ 1\ 1\ 0\ (14) \\
 \& 1\ 1\ 0\ 1\ (\sim 2) \\
 \hline
 1\ 1\ 0\ 0\ (12)
 \end{array}
 \quad
 \xrightarrow{\hspace{1cm}}
 \quad
 \begin{array}{r}
 1\ 1\ 0\ 0\ (12) \\
 \& 1\ 0\ 1\ 1\ (\sim 3) \\
 \hline
 1\ 0\ 0\ 0\ (8)
 \end{array}$$
  

$$\boxed{\begin{matrix} \times & 2 \\ 3 & 4 \end{matrix}}
 \quad
 \xrightarrow{\hspace{1cm}}
 \quad
 \boxed{\begin{matrix} \times & \times \\ 3 & 4 \end{matrix}}
 \quad
 \xrightarrow{\hspace{1cm}}
 \quad
 \boxed{\begin{matrix} \times & \times \\ \times & 4 \end{matrix}}$$

図 7: 値と値の反転との論理積の計算の例

## 4 性能の検証

本章では、復号アルゴリズムの性能を検証する。

### 4.1 検証方法

検証方法として3つの検証方法を考えた。

検証方法の1つ目は、ブロック誤りである。ブロック誤りとは、送信したラテン方陣と、復号して出来たラテン方陣を見比べ、ラテン方陣全体を一つのブロックとして復号できているかどうか見る。それに対して、雑音の強さとブロック誤り確率について検証する(図8)。

2つ目は、シンボル誤りである。シンボル誤りとは、送信したラテン方陣と復号して出来たラテン方陣の行列の各マスの数値を見比べる。また、各マスに対してのシンボル誤り確率を検証する(図9)。

3つ目は、復号成功の内訳である。送信したラテン方陣がそのまま送られる(パターン1)数と、復号したことで送信に成功した(パターン2)数をそれぞれの復号成功率で検証する(図10)。

プログラムの実装方法として、 $4 \times 4$  のラテン方陣の総数 576 個あるラテン方陣の中から、符号語としてランダムに1つを送信する。これを通信路に通す。ここで、図3の通信路を実装する際に、図11にプログラム用考案通信路を考えた。また、確率  $p$  で 1,2,3,4 が等確率に送られるような  $p$  を想定する。入力シンボルと出力シンボルが同じである確率を求めるとき、 $(1-p) + \frac{p}{4} = 1 - (\frac{3}{4}p)$  となる。理論上で入力シンボルと出力シンボルが同じである確率は、 $1-\epsilon$  である。また、プログラム上で  $p$  を設定するためには、 $p$  を  $\epsilon$  を使った式で表したい。等式を立てると、 $1 - (\frac{3}{4}p) = 1 - \epsilon$  となり、 $p = \frac{4}{3}\epsilon$  と設定した。

また、この通信路を用いてシミュレーション実験を行った。試行回数を100万回として、誤り確率を(復号誤り数)/(試行回数)で計測した。また、雑音の強さが限りなく小さいとき、3桁を超える失敗数を出すことができず、精度が落ちてしまうため、試行回数を1000万回、1億回といったように増やす場合もある。

1	2	3	4
2	1	4	3
3	4	2	1
4	3	1	2

復号成功  
=

1	2	3	4
2	1	4	3
3	4	2	1
4	3	1	2

1	2	3	4
2	1	4	3
3	4	2	1
4	3	1	2

復号失敗  
≠

2	3	4	1
1	2	3	4
3	4	1	2
4	1	2	3

図 8: ブロック誤り

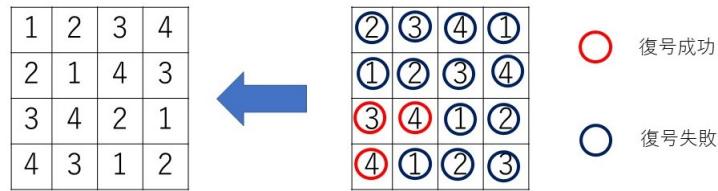


図 9: シンボル誤り

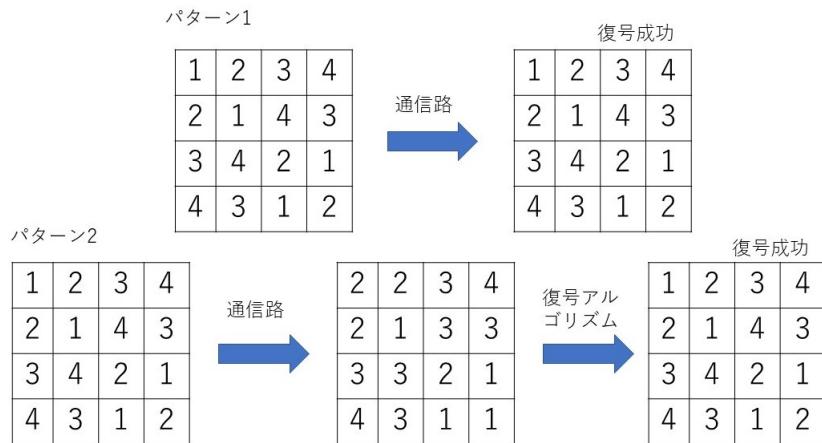


図 10: 復号成功の内訳

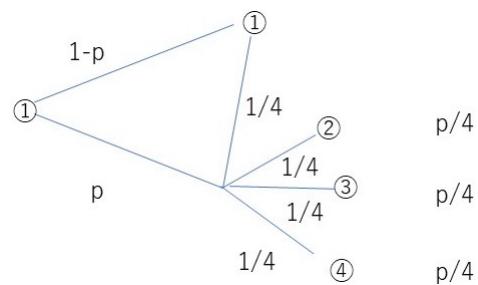


図 11: プログラム用考案通信路

## 4.2 結果と考察

雑音の強さを変えて、ブロック誤り確率とシンボル誤り確率について検証したグラフを図 12 に示す。このグラフから、ブロック誤り確率よりもシンボル誤り確率の方が雑音の強さに 対して、常に誤り確率が低いことがわかる。

次に、図 12 のグラフを対数スケールで表したものを見ると図 13 で示す。このグラフから、雑音の 強さが 0 になる前に復号誤り確率が 0 になる期待をしていたが、雑音の強さが 0 でないと、誤り確率が 0 にならなかった。

さらに、ラテン方陣の復号成功の内訳を図 14 に示す。グラフの見方としては、復号成功の中でも通信路を通ってきて、送信したラテン方陣のまま送られているものの割合をオレンジ色、通信路を通した符号語としてのラテン方陣に対して、復号アルゴリズムを適用して復号成功した割合を青色で示す。このグラフから、雑音の強さが弱いほど、元のラテン方陣がそのまま通ることが多くなる。また、元のラテン方陣が通らなかったとしても、雑音の強さが弱いときは、復号成功の数が多くなることが読み取れる。さらに、元のラテン方陣が雑音の影響を受け、別ものになったとしても、本研究の復号アルゴリズムを通して、元のラテン方陣を復号することが可能になったことがわかる。

最後に、先行研究結果の中で sum-product 復号を用いた研究がある [4]。先行研究 (左) と 本研究 (右) を比較したものを図 15 に示す。この図で用いた結果は、先行研究に合わせて、スケールを同じようにとったものとなっている。先行研究に比べ、シンプルなアルゴリズムではあるが、誤り確率は低くなかった。

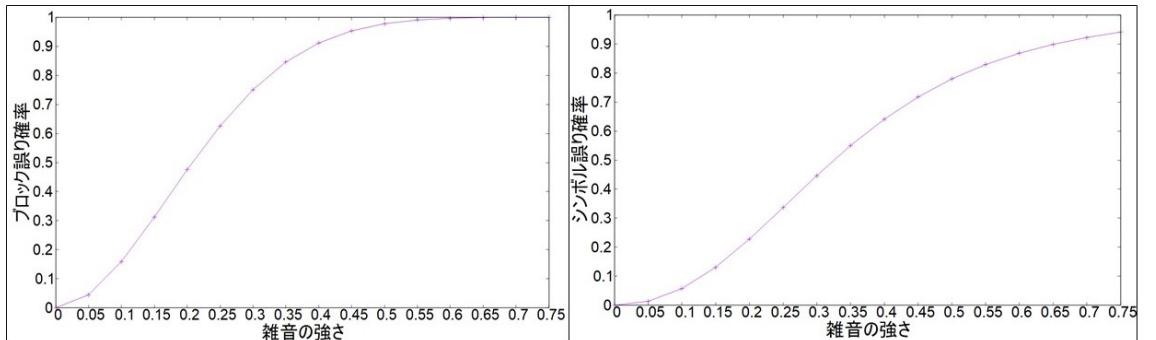


図 12: ブロック誤り確率とシンボル誤り確率

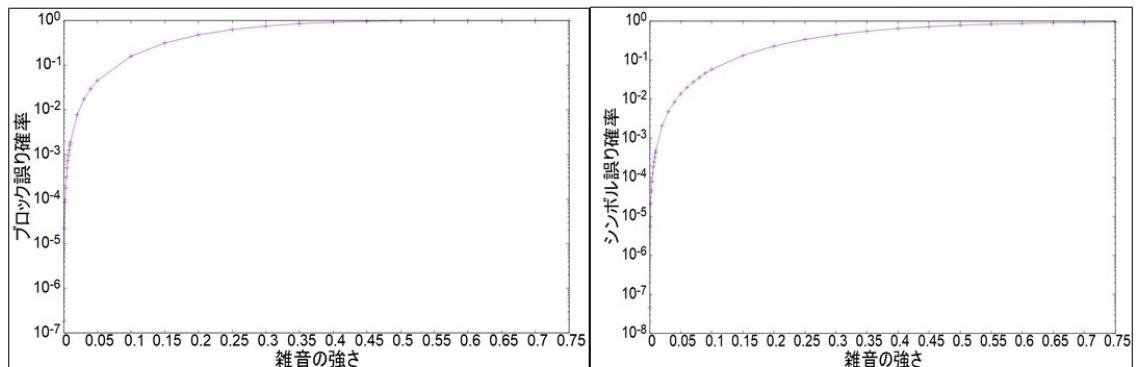


図 13: ブロック誤り確率とシンボル誤り確率 (log scale)

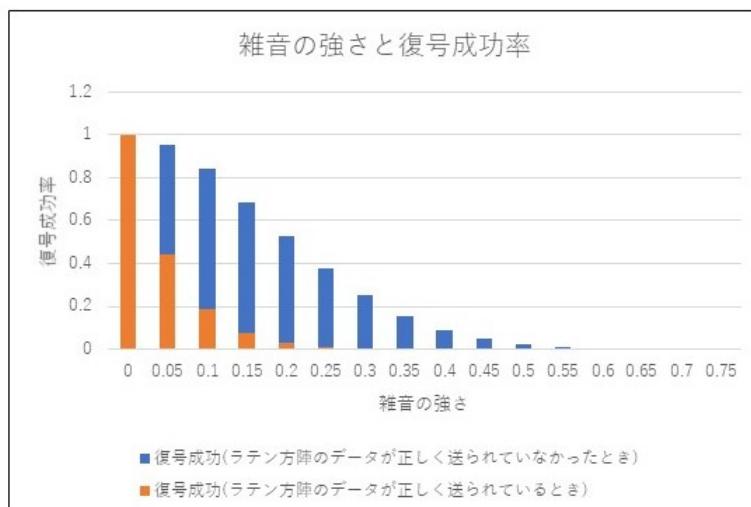


図 14: Latin 方陣の復号成功パターン

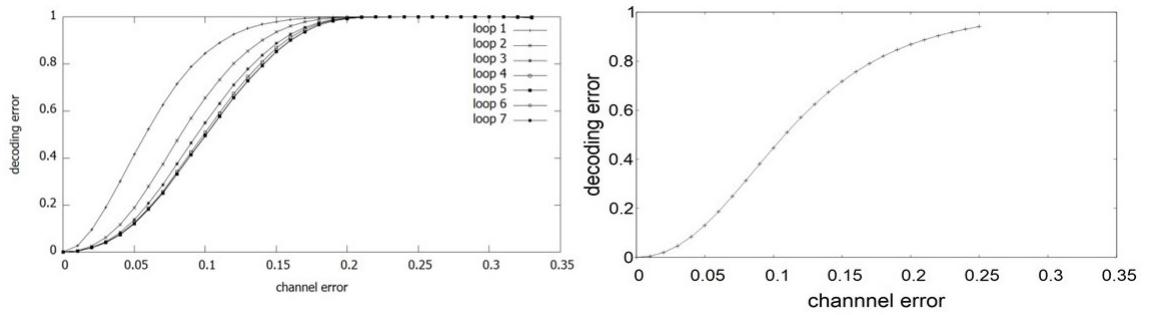


図 15: 先行研究 (左) と本研究 (右) の比較

## 5 まとめ

誤り訂正符号は通信の精度を高めるための技術であり、誤り訂正の技術は送りたい情報をあるルールを用いて符号化することによって行う。

本研究では、シンプルなアルゴリズムであったが、従来の復号方法を超えることができた。今後の展望としては、ラテン方陣の1マスに対してアルゴリズムを適用したとき、その1マスが変わることで、ほかのマスがどういう状態を持つのかパラメータ表示させ、それをアルゴリズムに組み込み、結果がよくなることを期待したい。それに加え、本研究の実験では通信路で雑音の発生源としてrand関数を用いた。ただ、rand関数は精度があまり高くないため、メルセンヌツイスター[6]という疑似乱数生成器を用いて、同様の実験でもより精度の高い検証を行いたい。

## 謝辞

本研究を行うにあたり、指導をしてくださった指導教員の西新幹彦准教授、また西新研究室の皆様に感謝の意を表する。

## 参考文献

- [1] 小林欣吾・森田啓義、情報理論講義、培風館、2008.
- [2] 韓太舜・小林欣吾、情報と符号化の数理、培風館、2006.
- [3] 比田井亮、「数独の消失通信路に対する復号誤り特性の測定について」、信州大学工学部卒業論文（指導教員：西新幹彦）、2012年3月。
- [4] 塚田芳寿、「誤り訂正符号としてのラテン方陣の復号方法について」、信州大学大学院総合理工学研究科修士論文（指導教員：西新幹彦）、2021年3月。
- [5] 岡田航、「ラテン方陣の効率の良い数え上げに向けた基礎的考察」、信州大学工学部卒業論文（指導教員：西新幹彦）、2018年3月。
- [6] Mersenne Twister:A random number generator (since 1997/10),  
<http://www.math.sci.hiroshima-u.ac.jp/m-mat/MT/mt.html>.2021年8月閲覧.
- [7] 長野透、「[C言語入門]乱数(rand)の使い方」、<https://www.sejuku.net/blog/25352>, 2021年8月閲覧.

## 付録 A ソースコード

#### A.1 ラテン方陣の復元成功, 失敗を数えるプログラム









```

        if (input2[i][l] == input2[i][j]) {
            input3[i][j] = 15;
        }
    }
}
for (int i = 0; i < NUM; i++) {
    for (int j = 0; j < NUM; j++) {
        for (int c = 0; c < NUM; c++) {
            if (input3[i][j] & (1 << c)) {
            }
            else {
            }
        }
    }
}
for (int i = 0; i < NUM; i++) {
    for (int j = 0; j < NUM; j++) {
        latin[i][j] = input3[i][j];
    }
}
touch = 1;
while (touch != 0) {
    touch = 0;
    for (int i = 0; i < NUM; i++) {
        for (int j = 0; j < NUM; j++) {
            if (latin[i][j] != 1 && latin[i][j] != 2 && latin[i][j] != 4 && latin[i][j] != 8) {
                for (int k = 0; k < 4; k++) {
                    if (k == i) {
                        continue;
                    }
                    if (latin[k][j] == 1 || latin[k][j] == 2 || latin[k][j] == 4 || latin[k][j] == 8) {
                        if (latin[i][j] & latin[k][j]) {
                            latin[i][j] = latin[i][j] & ~latin[k][j];
                            touch = 1;
                        }
                    }
                }
                for (int l = 0; l < 4; l++) {
                    if (l == j) {
                        continue;
                    }
                    if (latin[i][l] == 1 || latin[i][l] == 2 || latin[i][l] == 4 || latin[i][l] == 8) {
                        if (latin[i][j] & latin[i][l]) {
                            latin[i][j] = latin[i][j] & ~latin[i][l];
                            touch = 1;
                        }
                    }
                }
            }
        }
    }
}
for (int s = 0; s < NUM; s++) {
    for (int t = 0; t < NUM; t++) {
        if (latin[s][t] == 0) {
            x += 1;
        }
        else if (latin[s][t] != 1 && latin[s][t] != 2 && latin[s][t] != 4 && latin[s][t] != 8) {
            y += 1;
        }
    }
}
if (x == 0 && y == 0) {
    x = y = 0;
}
else if (x != 0 && y == 0) {
    x = y = 0;
}

```

```

    else if (x == 0 && y != 0) {
        x = y = 0;
    }
    else if (x != 0 && y != 0) {
        x = y = 0;
    }
    for (int i = 0; i < NUM; i++) {
        for (int j = 0; j < NUM; j++) {
            for (int c = 0; c < NUM; c++) {
                if (latin[i][j] & (1 << c)) {
                }
                else {
                }
            }
        }
    }
}
int flag = 0;
for (int i = 0; i < NUM; i++) {
    for (int j = 0; j < NUM; j++) {
        input1[i][j] = 1 << (input1[i][j] - 1);
        if (input1[i][j] == latin[i][j]) {
            continue;
        }
        else {
            same -= 1;
            differ += 1;
            flag = 1;
            break;
        }
    }
    if (flag == 1) {
        break;
    }
    same += 1;
}
printf("same = %d\n", same);
printf("differ = %d\n", differ);
return(0);
}

```