

Boolean Groebner 基底を用いた 数独パズルの数学的難易度指標の相関について

中野 哲夫
進藤 未来
吉原 元

東京電機大学大学院理工学研究科

12月19日(月)

目次

- 1 概要
- 2 準備
- 3 井上アルゴリズムと CII アルゴリズム
- 4 戦術閉包と s_{∞} -rank
- 5 主要結果
- 6 参考文献

目次

- 1 概要
- 2 準備
- 3 井上アルゴリズムと CII アルゴリズム
- 4 戦術閉包と s_{∞} -rank
- 5 主要結果
- 6 参考文献

概要

ブーリアングレブナ基底を用いて数独パズルを解く数学的解法として、井上アルゴリズム [4] がある。井上アルゴリズムは人間のトライアンドエラー方式の解法と似ているため、数独パズルの難易度判定に用いられて成果を上げている。

井上アルゴリズムに基づく数独パズルの数学的難易度指標には、大きく分けて2つある。すなわち、

(1) 佐藤先生、井上先生のグループによる **b-rank**, **s-rank**[11] およびそれらから派生する rank (s_{∞} -rank など)

(2) 講演者達のグループが提唱した CII アルゴリズムを用いる指標 **CII**[6]-[11] (Computable Inoue Invariant, 計算可能井上不変量) と **SMY 不変量** [15][16] (Shindou-Mikoshiba-Yoshihara 不変量, SMYI)

の2種類である。

- 本発表では、これら2種類の数学的指標の相関について、実験的に検証した結果を報告する。より詳しく、超難問について、CII アルゴリズムに基づいて定義される**不変量 SMYI**と、**MDSL**[15][16]、 **s_∞ -rank** (s_∞ -rank 有限の場合)、**LAC** (s_∞ -rank 無限の場合)のそれぞれとの相関について実験的に検証した結果を報告する。特に、 **s_∞ -rank 無限の場合、SMYI と LAC は強い負の相関がある**という興味深い結果を得た。
- 最も難しい s_∞ -rank 無限のパズルの例は従来3つしか知られていなかったため、このクラスのパズルの階層付けはほとんど研究されていなかった。しかし、今回、約150問の s_∞ -rank 無限のパズルの例が新たに見つかり、実験により **LAC がこのクラスの階層付けの良い指標**になることが検証できた。
- 本研究では、数独の実験でよく用いられる郷内氏の7冊の標準的問題集(735問)[2]および、インカラ博士の作成した超難問12問[3]、ターレック氏の超難問編634問[12]を実験サンプルとして用いた。ターレック氏の634問は今回新たにMagma形式の入力で採録したものである。
- 計算に用いるプログラムは、数式処理システム **Magma**[1]上で開発・実装した。

目次

- 1 概要
- 2 準備
- 3 井上アルゴリズムとCIIアルゴリズム
- 4 戦術閉包と s_∞ -rank
- 5 主要結果
- 6 参考文献

ブリアングレブナ基底の定義と基本性質

定義1 (ブール環)

A を単位元をもつ可換環とする。任意の $a \in A$ について $a^2 = a$ が成立するとき、 A を **ブール環**という。

$\mathbb{F}_2 = \{0, 1\}$ を2元体、 \mathbb{F}_2^m を成分ごとの加法、乗法について可換環とみなすと、 \mathbb{F}_2^m はブール環となる。次に、 \mathbb{F}_2^m に係数をもつブール多項式環を定義しよう。

定義2 (ブール多項式環)

$\mathbb{F}_2^m[x] = \mathbb{F}_2^m[x_1, \dots, x_n]$ を \mathbb{F}_2^m に係数にもつ n 変数多項式環とする。 $\mathbb{F}_2^m[x]$ 自体は、ブール環ではない。そこで

$$\mathbb{F}_2^m(x) := \mathbb{F}_2^m[x_1, \dots, x_n] / \langle x_1^2 - x_1, \dots, x_n^2 - x_n \rangle.$$

とおくと、 $\mathbb{F}_2^m(x)$ はブール環となる。これを **ブール多項式環**と呼ぶ。

次に、まず、 $\mathbb{F}_2^m[x] = \mathbb{F}_2^m[x_1, \dots, x_n]$ のグレブナ基底を説明する。 $\mathbb{F}_2^m[x]$ でも、通常のを係数とする多項式と同様に、余りのある割り算アルゴリズムが(少しの修正をした上で)成立する(余りの条件が少し異なる)。

定義3 (グレブナ基底)

$\mathbb{F}_2^m[x]$ 上に単項式順序を1つ与え、 $I \subset \mathbb{F}_2^m[x]$ をイデアルとする。 $G := \{g_1, \dots, g_s\}$ を I の有限部分集合とし、 $\text{LT}(I) = \{\text{LT}(f) \mid f \in I, f \neq 0\}$ を I の元の先頭項の集合とする。このとき、

$$\langle \text{LT}(I) \rangle = \langle \text{LT}(g_1), \dots, \text{LT}(g_s) \rangle$$

が成り立つならば、 G は I の **グレブナ基底**という。

定義4 (極小グレブナ基底, 簡約グレブナ基底)

- (1) G をイデアル I のグレブナ基底とする。任意の $g \in G$ について、 $\text{LT}(g) \notin \langle \text{LT}(G \setminus \{g\}) \rangle$ が成立するとき、 G を **極小グレブナ基底**という。
- (2) G を極小グレブナ基底とする。任意の $g \in G$ に対して、 $\bar{g}^{G \setminus \{g\}} = g$ が成り立つとき、 G を **簡約グレブナ基底**という。ここで、 $\bar{g}^{G \setminus \{g\}}$ は、 g を $G \setminus \{g\}$ で割った余りである。

体上の多項式環の場合と違って、簡約グレブナ基底は一意に定まらないので、次の階層的グレブナ基底が必要である。

定義 5 (階層的グレブナ基底)

G をイデアル $I \subset \mathbb{F}_2^m[x]$ の簡約グレブナ基底とする。もし任意の $f, g \in G (f \neq g)$ について $\text{LM}(f) \neq \text{LM}(g)$ ならば、 G は階層的グレブナ基底と呼ばれる。

任意のイデアル $I \subset \mathbb{F}_2^m[x]$ は、唯一の階層的グレブナ基底をもつ。これは体上の係数の場合と似た（しかし、少し異なる）Buchberger アルゴリズムによって証明できる。

以上の $\mathbb{F}_2^m[x]$ のイデアルのグレブナ基底の定義および性質は、ブール多項式環 $\mathbb{F}_2^m(x)$ のイデアルについても全く同様に成立する。

例えば、

定義 6 (ブーリアングレブナ基底)

$\mathbb{F}_2^m[x]$ に単項式順序を 1 つ与えると、自然に、 $\mathbb{F}_2^m(x)$ の（標準単項式の集合に）単項式順序が誘導される。 I を $\mathbb{F}_2^m(x)$ のイデアルとし、 $G := \{g_1, \dots, g_s\}$ を I の有限部分集合とする。このとき、 $\langle \text{LT}(I) \rangle = \langle \text{LT}(g_1), \dots, \text{LT}(g_s) \rangle$ が成り立つならば、 G を I のブーリアングレブナ基底という。

階層的グレブナ基底の存在と一意性も、ブール多項式環で成立する。ブール多項式環のイデアルの具体的な計算は、すべてブーリアングレブナ基底を用いて遂行される。

数独パズルの数学的定式化

次に数独パズルのルールをブール多項式環を用いて定式化する。 x_{ij} を i 行 j 列のマスに対応する変数とし、 \mathbb{F}_2^9 上の 81 変数 x_{11}, \dots, x_{99} のブール多項式環

$$\mathbb{F}_2^9(x) = \mathbb{F}_2^9(x_{11}, \dots, x_{99})$$

を考える。 $e_1 := (1, 0, \dots, 0), \dots, e_9 := (0, \dots, 0, 1), 0 = (0, \dots, 0), 1 = (1, \dots, 1) \in \mathbb{F}_2^9$ とおく。各 e_i は、数独のマス埋める $\{1, 2, \dots, 9\}$ の値 i に相当する。マス x_{ij} の初期値（ヒント）が $k (1 \leq k \leq 9)$ の場合は、方程式 $x_{ij} - e_k = 0$ でこの条件が表される。

数独パズルの一行目に注目すると、パズルのルールは次の 37 本のブール方程式で表せる。

$$\sum_{i=1}^9 x_{1i} = 1, x_{1i} \cdot x_{1j} = 0 (i \neq j).$$

数独パズルには 9 行、9 列、9 個の 3×3 のブロックがあるため、合計で $37 \times 27 = 999$ 本のブール方程式でパズルのルールは表せる。 $I \subset \mathbb{F}_2^9(x)$ を初期値と 999 本のブール多項式によって生成されるイデアルとする。このとき、 I を与えられた数独パズルの初期イデアルという。

定義 7 (2 種類の解集合)

イデアル $I \subset (\mathbb{F}_2)^9(x)$ に対して

$$\mathbb{V}(I) := \{(a_{11}, \dots, a_{99}) \mid a_{ij} \in \mathbb{F}_2^9, \text{すべての } f \in I \text{ について}$$

$$f(a_{11}, \dots, a_{99}) = 0\} \subset (\mathbb{F}_2^9)^{81}$$

$$\mathbb{VS}(I) := \{(a_{11}, \dots, a_{99}) \mid a_{ij} \in \{e_1, \dots, e_9\}, \text{すべての } f \in I \text{ について}$$

$$f(a_{11}, \dots, a_{99}) = 0\}$$

と定義する。 $\mathbb{V}(I)$ はイデアル I の定めるアフィン多様体であり、一方、 $\mathbb{VS}(I)$ は各変数が e_1, \dots, e_9 のみを値として取ることができる解集合である。したがって、 I が数独パズルのイデアルの場合、 $\mathbb{VS}(I)$ のほうを用いる。

定義 8 (極小多項式)

$I \subset \mathbb{F}_2^9(x)$ が $\mathbb{V}(I) \neq \emptyset$ を満たすようなイデアルとする。1 つの変数 x_i に対して、イデアル $I \cap \mathbb{F}_2^9(x_i)$ は $\mathbb{F}_2^9(x_i)$ の単項イデアルであり、次のように書ける。 $h(x_i) \in (\mathbb{F}_2^9)^9(x_i)$ に対して、

$$I \cap (\mathbb{F}_2^9)^9(x_i) = \langle h(x_i) \rangle.$$

この 1 変数多項式 $h(x_i) = ax_i + b (a, b \in \mathbb{F}_2^9)$ は I と x_i によって一意的に決まり、変数 x_i に対する I の極小多項式という。

極小多項式は、変数 x_i の値に関する重要な情報を持っている（下記の命題 10 参照）。

定義 9 (1 変数多項式の分類)

$f(x_j) = ax_j + b \in \mathbb{F}_2^9(a, b \in \mathbb{F}_2^9)$ を変数 x_j の 0 でない 1 変数ブール多項式とする。

(1) **解多項式** (solution-polynomial)

ある k に対して, $f = x_j + e_k$ の形をしている場合, 解多項式と呼ぶ。

(2) **準解多項式** (semisolution-polynomial)

$a \neq 1, a \cdot e_k \neq 0$ である $a \in \mathbb{F}_2^9$ に対して, $f = a \cdot x_j + e_k$ の形をしている場合, a 型の準解多項式と呼ぶ。

$\#a = 8$ の $a \in \mathbb{F}_2^9$ に対して, $f = a \cdot x_j$ の形をしている場合, b 型の準解多項式と呼ぶ。ここで, $\#a$ は a に含まれる 1 の数のことである。ある k に対して, $a = 1 + e_k$ と表されることに注意する。

(a) 型, (b) 型ともに, 解多項式 $g = x_j + e_k$ は, この準解多項式 f に対応する解多項式と呼ばれ,

$\text{Asp}(f)$ で表す。

(3) **矛盾多項式** (contradiction-polynomial)

f が次のいずれかの形をしている場合, f を矛盾多項式と呼ぶ。

(a) $f = b(b \in \mathbb{F}_2^9, b \neq 0)$.

(b) $f = x_j$.

(c) $f = ax_j + b(a \neq 0, \#b \geq 2)$.

(d) $f = ax_j + e_k(a \neq 0, 1, a \cdot e_k = 0)$.

(4) **分岐多項式** (branch-polynomial)

$f = a \cdot x_j (1 \leq \#a \leq 7)$ の形をしている場合, f を分岐多項式と呼ぶ。また, $\#a$ をこの分岐多項式の長さと呼ぶ。

次の命題は, 極小多項式がもつ変数の情報を示す。

命題 10

f を変数 x_j に関するイデアル I の極小多項式とする。このとき, 次が成り立つ:

(1) もし f が解多項式 $f(x_j) = x_j + e_k$ ならば $x_j + e_k = 0$, つまり $x_j = e_k$ となる。したがって, 変数 x_j の値は k と決まる。

(2) もし f が $\text{Asp}(f) = x_j + e_k$ に対応する準解多項式であるとき, 変数 x_j の値は $\text{VS}(I)$ 内で e_k と決まる。したがって, 準解多項式は解多項式ではないが, 変数 x_j の値を一意に決定できる。

(3) $f = ax_j (1 \leq \#a \leq 7)$ を分岐多項式とする。このとき,

$a = 1 + e_{i_1} + \dots + e_{i_p} (1 \leq i_1 < \dots < i_p \leq 9, 2 \leq p = 9 - \#a \leq 8)$ と書くと, 方程式 $f = 0$ は $\text{VS}(I)$ 内で p 個の解 e_{i_1}, \dots, e_{i_p} を持つ。したがって, 変数 x_j の取り得る値は $\{e_{i_1}, \dots, e_{i_p}\}$ と分かる。

次に, 井上アルゴリズムと CII アルゴリズムを実行する際に重要である, 基本閉包を定義する。

定義 11 (基本閉包)

(1) $I \subset \mathbb{F}_2^9(x)$ をイデアルとする。 $\text{SemiSol}(I)$ を I の準解多項式の集合, $\text{Asp}(\text{SemiSol}(I))$ を $\text{SemiSol}(I)$ の準解多項式に対応する解多項式の集合とする。ここで,

$$\Psi_0(I) := I + \langle \text{Asp}(\text{SemiSol}(I)) \rangle$$

と定義する。 $\Psi_0(I)$ は I の準解多項式に対応する解多項式を I にすべて付け加えて拡張されたイデアルになる。

(2) 次のようなイデアルの昇鎖列を考える。

$$I \subset \Psi_0(I) \subset (\Psi_0)^2(I) \subset \dots \subset \mathbb{F}_2^9(x).$$

$\mathbb{F}_2^9(x)$ は有限集合なので, $(\Psi_0)^m(I) = (\Psi_0)^{m+1}(I) = \dots$ を満たすような m が存在する。そのような m に対して, $\text{BC}(I) := (\Psi_0)^m(I)$ と定義し, I の**基本閉包**と呼ぶ。

定義 12 (極大イデアル)

I を各変数の 81 個の解多項式によって生成されるイデアル $I := \langle x_1 + e_{i_1}, \dots, x_{81} + e_{i_{81}} \rangle$ とする。このとき, I を**極大イデアル**と呼ぶ。もし, I が極大イデアルならば, $\text{VS}(I) = \{e_{i_1}, \dots, e_{i_{81}}\}$ となる。したがって, 極大イデアルはすべてのマス (変数) の値が決まった, 解けた形のイデアルである。

定義 13 (基本可解と基本ランク)

$I \subset \mathbb{F}_2^9(x)$ を $\text{BC}(I)$ が極大イデアルであるようなイデアルとする。このとき, I を**基本可解なイデアル**と呼ぶ。もし I が数独パズルのイデアルならば, I は**基本可解なパズル**という。また, I を基本可解とするとき, $(\Psi_0)^m(I)$ が極大イデアルとなるような最小の m をイデアル I の**基本ランク**という。 I が基本可解ではない場合, I の基本ランクは ∞ と定める。

初級から中級レベルのパズルは基本可解であることが知られており, また, 基本ランクは基本可解なパズルに対する難易度判定の優れた指標であることがわかっている。

アルゴリズム 14 (井上アルゴリズム)

入力：与えられた数独パズルの初期イデアル I_0 . I_0 は井上樹形図のルートである.

出力：解のデータ L . L の初期値は空配列とし、解が見つかるたびに解のデータを L に格納していく.

(1) 初期イデアル I_0 の基本閉包をとり、このイデアルを $J := BC(I)$ とする. J は、井上樹形図における深さ 1 の節である.

(2) ここで 3 つの場合がある.

(a) J が矛盾多項式を含む場合、 L には何も格納しない. このとき、井上樹形図において、 J は非解葉と呼ばれる終点となる.

(b) J が矛盾多項式を含まずに、極大イデアルとなる (J が 81 個の解多項式で生成される) 場合、 L には解のデータ (81 個のマス値) を格納する. このとき、井上樹形図において、 J は解葉と呼ばれる終点となる.

(c) J が非解葉でも解葉でもない場合、 J は分岐点と呼ばれる. このとき、任意の空きマス x_{ij} を選び、その取り得る値が e_{k_1}, \dots, e_{k_p} の p 個あるとする. 各 $l (1 \leq l \leq p)$ に対して、 $I := I + \langle x_{ij} + e_{k_l} \rangle$ と再定義し、(1) に戻る.

(3) この操作をすべての節が終点になるまで繰り返す. このアルゴリズムは有限回で終了し、その結果解のデータ L と井上樹形図が得られる.

MDSL (Minimal Depth of Solution Leaves, 解葉の最小深さ)

- 井上アルゴリズムを実行すると、分岐点で任意の空きマスを選んで良いため、多くの井上樹形図があり得る. それら全ての井上樹形図を考え、各井上樹形図の解葉の深さの最小値 (the Minimum Depth of Solution Leaves) を **MDSL** という.
- 郷内氏の 735 問、インカラ博士の 12 問、ターレック氏の 634 問合計 1381 問の MDSL を計算した結果、すべての問題で **MDSL=3 以下** となった. さらに、現在のところ MDSL が 4 以上のパズルは見つかっていない.
- MDSL=3 であるパズルは超難問**に限られることがわかっている (後の関連の部分を参照).

そこで講演者達は次の予想を提唱している.

MDSL 予想 [15],[16]

井上アルゴリズム (トライアンドエラー方式) で数独を解くとき、**唯一解をもつ数独パズルの解葉の最小の深さは必ず 3 以下**であろう.

これは、どんなに難しいパズルでも分岐点で 2 回適当な空きマスと値を選べば、必ず解に到達するというとても強い予想である.

CII アルゴリズム

定義 15 (理想的井上不変量)

$I \subset \mathbb{F}_2^9(x)$ を数独パズルの初期イデアルとする. 井上アルゴリズムから生じる全ての井上樹形図を考慮し、それらの井上数の最小値をイデアル I (または数独パズル) の**理想的井上不変量** (Virtual Inoue Invariant, **VII**) と呼ぶ. ただし、2 つの井上数は、左から比較する辞書式順序で大小を比較することにする.

VII は数独の不変量ではあるが、実際には井上樹形図の数が多すぎて計算不可能.

定義 16 (CII アルゴリズム, 計算可能井上不変量 (CII))

$I \subset \mathbb{F}_2^9(x)$ を数独パズルの初期イデアルとする. **CII アルゴリズム**では、分岐点で任意の空きマスを選ぶのではなく、次のように**できるだけ井上樹形図が小さくなりそうな変数に絞り**、**それらの変数のみを選択**できることにする. 最適変数絞り (A)-(D) または (0) を 1 つ固定し、各分岐点で指定した絞り方によって変数を選ぶことにする. このとき、生成されるすべての井上樹形図における井上数の最小値を I の**計算可能井上不変量** (Computable Inoue Invariant, **CII**) と呼ぶ. CII は、分岐点での最適変数の絞り方に依存するが、絞り方を 1 つ決めてしまえば、立派なイデアルの数学的不変量となる. 選んだ最適変数絞りを明確にするために、CII.D, CII.C などと表す. ここで、CII.0 は VII と同じである. CII アルゴリズムの実行結果は CII 樹形図で簡明に表すことができる.

変数絞りのための 5 つの基本戦術

(A) この分岐点で分岐数最小の変数に絞る.

(B) まず A で空きマスを絞ったあとに、それらの取り得る値を入れてできる次の節たちでの分岐数 (の和) が最小のマスに絞る. ただし、葉の分岐数は 0 と約束する.

(C) B まで絞ったあとに、次の節たちでの空きマスの個数 (の和) が最小の空きマスに絞る.

(D) C で絞ったあとに、次の節たちでの分岐数最小の空きマスの個数 (の和) が最小の空きマスに絞る.

(0) 便宜上、分岐点で全く変数を絞らずに、全ての空きマスを選んで良い場合も、CII アルゴリズムの 1 つの場合に含め、この場合は (0) (絞りが無いという意味) と表すことにする. この場合は、本質的には井上アルゴリズムと同じである.

一番強い D 絞りで、多くの場合、各分岐点で選択できる変数 (空きマス) の個数を数個程度まで絞ることができる.

図 2: CII 樹形図 (L15) の例

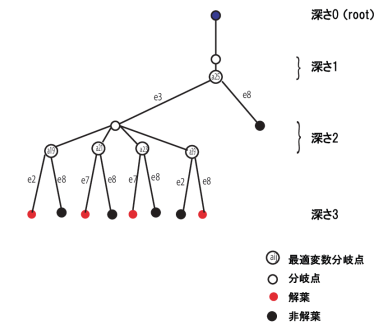
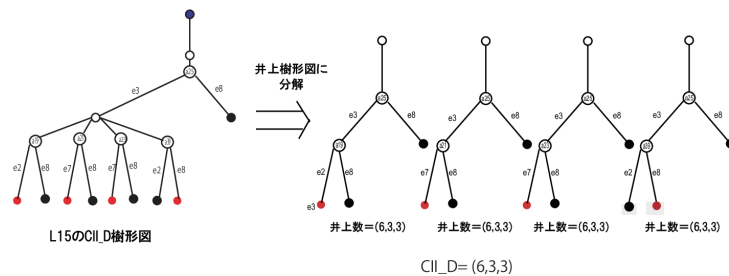


図 3: CII-D 樹形図 (L15) を井上樹形図に分解して CII-D を求める.



不変量 SMYI (Shindou-Mikoshiba-Yoshihara Invariant)

CII は数独パズルの良い難易度指標であることが分かっているが、超難問については、まれに、計算量が多く実行に時間がかかりすぎて計算できないことがある. そこで、CII 樹形図の大きさを直接はかる不変量 SMYI を次のように定義する.

定義 17 (SMYI(Shindou-Mikoshiba-Yoshihara Invariant))

T を数独パズルの CII 樹形図とする. $a := T$ のすべての節の数 (2 種類の分岐点については、最適変数分岐点を数え、通常分岐点を数えない.), $b := T$ の葉の数, $c := T$ の深さの最大値とし、 (a, b, c) の 3 つ組を **SMY 不変量** (略して **SMYI**) と定義する.

- SMYI は分岐点での変数の絞り方に依存するが、一度変数の絞り方を決めてしまえば、数独 (イデアル) の不変量である.
- そこで、**最適変数絞り (D)** を採用し、これを SMYI として郷内氏の問題集 (全部で 735 問) を使用した実験を行った.

- SMYI の欄の値は、節の個数の平均値を表している。また、郷内レベルとは郷内氏の割り振った数独の難易度レベルのことである。
- SMYI は、郷内レベルと約 **0.92** という高い正の相関関係をもつことが判明した。これにより、SMYI は数独パズルの難易度の良い数学的指標として使用できることが検証された。

表 1 : SMYI の郷内レベルとの相関

郷内レベル	1	2	3	4	5	6	7	8	9	10
SMYI	2.00	2.00	2.00	2.00	2.00	2.00	2.00	14.11	14.28	23.03
郷内レベル	11	12	13	14	15	16	17	18		相関係数
SMYI	26.79	29.01	20.38	43.79	47.45	29.17	37.40	57.60		0.91618

目次

- 概要
- 準備
- 井上アルゴリズムとCIIアルゴリズム
- 戦術閉包と s_{∞} -rank
- 主要結果
- 参考文献

戦術閉包の定義

- 初期イデアルにおいて、任意の空きマス x_j を選び、1-9 までの値 e_k を順に試し、一意的に決まるマスをすべて埋めていって途中で矛盾が起きた場合、 x_j には e_k が入らないという情報が得られる。この情報を全部初期イデアルに付け加えてから、基本閉包をとる操作を考える。
- この操作を何度もやると初期イデアルが拡張される
→ イデアルの拡張ができなくなったときのイデアル = **戦術閉包**

戦術閉包を取る操作は、人間の解法テクニックのいくつかを含む (XYwing 法, XY-chain 法など) [5], 非常に強力な数独の数学的解法であり、戦術閉包をとっても解けないパズル (戦術閉包が極大イデアルにならないパズル) は (講演者達の知る限り) これまで3例しか知られていなかった。

定義 18 (戦術閉包)

・ $I \subset \mathbb{F}_2^9(x)$ をイデアルとする。解多項式 $x_j + e_i$ が **rank k の反駁解多項式** とは、イデアル $\Psi_0^k(I + \langle x_j + e_i \rangle)$ が矛盾多項式を含むことをいう。ただし、 k はそのような最小値を意味する。
 ・ 自然数 k に対して、 $BR_k(I) := \{e_i x_j \mid x_j + e_i \text{ が } I \text{ の rank } k \text{ 以下の反駁解多項式}\}$ とする。また、 $\Phi_k(I) := \text{BC}(I + \langle BR_k(I) \rangle)$ と定義すると、イデアルの昇鎖列

$$I \subset \Phi_k(I) \subset (\Phi_k)^2(I) \subset \dots$$

ができる。 $\mathbb{F}_2^9(x)$ は有限集合であるから、ある自然数 n が存在して $(\Phi_k)^n(I) = (\Phi_k)^{n+1}(I) = (\Phi_k)^{n+2}(I) = \dots$ となる。この n について、 $\Phi_k^*(I) := (\Phi_k)^n(I)$ と定義して、これを **rank k の I の戦術閉包** という。次に k を動かすと次のイデアルの昇鎖列

$$\Phi_1^*(I) \subset \Phi_2^*(I) \subset \dots \subset \Phi_k^*(I) \subset \Phi_{k+1}^*(I) \subset \dots$$

ができるが、 $\mathbb{F}_2^9(x)$ は有限集合であるから、ある自然数 l が存在して $\Phi_l^*(I) = \Phi_{l+1}^*(I) = \Phi_{l+2}^*(I) = \dots$ となる。この l について、 $\Phi_{\infty}^*(I) := \Phi_l^*(I)$ と定義して、これを I の **戦術閉包** という。

定義 19 (s-rank)

$I \subset \mathbb{F}_2^q(x)$ を戦術閉包が極大イデアルになる (戦術可解イデアルという) ようなイデアルとして, $\Phi_k^*(I) = \Phi_\infty^*(I)$ となる最小の k をとる. この k を I の **s-rank** という. ただし, 基本可解イデアルの s-rank は 0 とする. また, $\Phi_\infty^*(I)$ が極大イデアルでないとき, s-rank は ∞ とする.

戦術閉包 (と s-rank) を実際に計算するには, 上記の定義は, 昇鎖を 2 回取る必要があり, また, 昇鎖が一旦止まってもその後も止まり続けるとは限らないため, 定義通りの計算はやや面倒になる. そこで, この発表では次の戦術閉包の定義および s_∞ -rank を用いることにする.

定義 20 (戦術閉包の定義 2 と s_∞ -rank)

$I \subset \mathbb{F}_2^q(x)$ を基本可解でないイデアルとする. 解多項式 $x_j + e_k$ が反駁解多項式とは, イデアル $BC(I + \langle x_j + e_k \rangle)$ が矛盾多項式を含むことをいう.

$$BR_\infty(I) := \{e_k x_j \mid x_j + e_k \text{ は } I \text{ の反駁解多項式}\}$$

とする. さらに, $\Phi_\infty(I) := BC(I + \langle BR_\infty(I) \rangle)$ とすると, 次の昇鎖列

$$I \subset \Phi_\infty(I) \subset (\Phi_\infty)^2(I) \subset \dots$$

ができる. この昇鎖はある自然数 n 以後は安定し, $(\Phi_\infty)^n(I) = (\Phi_\infty)^{n+1}(I) = \dots$ となる. このとき, $\Phi_\infty^*(I) = (\Phi_\infty)^n(I)$ (戦術閉包) となる. また, これを満たす最小の n を I の **s_∞ -rank** という. ただし, 基本可解イデアルの s_∞ -rank は 0 とする. また, 戦術閉包 $\Phi_\infty^*(I)$ が極大イデアルでないとき, s_∞ -rank は ∞ とする. s_∞ -rank が ∞ のイデアルについては, 上の最小の n を **昇鎖の長さ** (LAC, the Length of Ascending Chain) と呼ぶことにする.

s-rank が ∞ であることと, s_∞ -rank が ∞ であることは同値であり, また, s-rank が有限値のパズルについては, s_∞ -rank は s-rank より少し値が小さくなる.

次の表は, ターレック氏の 634 問に含まれる s_∞ -rank が ∞ のパズル (158 問) についての, LAC の計算結果である. なお, これまでに講演者達の知る限りでは s_∞ -rank = ∞ であるパズルは 3 例しか知られていなかった.

表 2: ターレック氏の s_∞ -rank ∞ のパズルの内分け

昇鎖の長さ	0	1	2	3	4
B (ターレック氏)	5	9	0	0	0
HD (ターレック氏)	67	60	8	5	3
SM (ターレック氏)	0	0	1	0	0
問題数	72	69	9	5	3

目次

- 1 概要
- 2 準備
- 3 井上アルゴリズムと CII アルゴリズム
- 4 戦術閉包と s_∞ -rank
- 5 主要結果
- 6 参考文献

SMYI と MDSL, s_∞ -rank, LAC の相関

- 郷内氏の標準的な 735 問については、人間の難易度として「郷内ランク」がついていたので、これが人間の難易度の目安となった。さらに、この 735 問では CII を計算するのにそれほど時間がかからず、CII, SMYI とともに郷内ランクと高い相関があった。
- 一方、ターレック氏の超難問 634 問には、人間の難易度が付与されていないため、難易度の目安となる指標がない。そこで、超難問の場合、ここでは SMYI を難易度の目安として採用することにする。

【SMYI を難易度の目安として採用する根拠】

- CII アルゴリズムは、人間のトライアンドエラー方式の解法を模しているため、CII アルゴリズムの実行結果の大きさを直接はかる SMYI は人間の難易度に近いと考えられること。
- 郷内氏の 735 問の問題では、郷内ランクと SMYI は高い相関があったこと。
- CII, SMYI とともに郷内ランクと高い正の相関があり、どちらも難易度の目安として使えるが、CII は超難問については計算に時間がかかりすぎるため、全部計算するには無理がある。一方、SMYI は超難問でも 1 問数分程度で計算できること。

- そこで、超難問については、SMYI を難易度の指標と考え、実験で
 - SMYI と MDSL
 - SMYI と s_∞ -rank (s_∞ -rank 有限の場合)
 - SMYI と昇鎖の長さ (the Length of Ascending Chain, LAC) (s_∞ -rank 無限の場合)
 の比較を行い、傾向が判明したので報告する。
- SMYI の欄の値は、節の個数の平均値を表している。

(1) SMYI と MDSL

- SMYI と MDSL は 0.98 という高い正の相関関係をもつことが判明した。これは、MDSL が大きいほど SMYI の値は大きくなり、パズルが難しくなることを表している。

表 3 : SMYI と MDSL

MDSL	1	2	3	
SMYI	2.00	213.37	603.02	相関係数
問題数	11	270	365	0.98565

(2) SMYI と s_{∞} -rank (s_{∞} -rank 有限の場合)

- s_{∞} -rank が 1-5 までの場合は、0.95 という高い正の相関関係をもつことが判明した。これは、 s_{∞} -rank が大きいほど SMYI の値は大きくなり、パズルが難しくなることを表している。
- s_{∞} -rank が 1-6 までの場合は、 s_{∞} -rank=6 である問題数が 2 問しかなくどちらも SMYI が小さいため、相関係数が低くなってしまふ。

表 4 : SMYI と s_{∞} -rank

s_{∞} -rank	1	2	3	4	5	6		
SMYI	183.01	288.69	359.88	661.87	1029.75	254.00	相関係数	
問題数	168	247	49	15	4	2	0.9517641	0.4756188

(3) SMYI と昇鎖の長さ (LAC) (s_{∞} -rank 無限の場合)

- SMYI と LAC は -0.96 という高い負の相関関係をもつことが判明した。これは、LAC が大きいほど SMYI の値は小さくなり、パズルが簡単になる、また、LAC が小さいほど SMYI の値は大きくなり、パズルが難しくなることを表している。

表 5 : SMYI と LAC

昇鎖の長さ	0	1	2	3	4	
SMYI	917.26	931.31	771.22	610.67	425.75	相関係数
問題数	72	70	9	6	4	-0.96083

- 発表者達は、当初、 s_{∞} -rank が有限値の場合と同様に、LAC が大きいほどパズルは難しいのではないかと考えていたので、これはとても意外な結果であった。このようになる直観的な理由としては、例えば LAC=0 の場合は、戦術閉包をとっても全く空きマスが埋まらない (イデアルが拡張されない) ので、戦術閉包という強力な解法に対して抵抗力が最大であると考えられる。よって、LAC が小さいほど難しい、という結果になったと解釈できる。
- この結果により、LAC が s_{∞} -rank ∞ のパズルの階層付けの良い指標になることが実験的に検証できた。
- 以上の結果より、 s_{∞} -rank= ∞ のパズルで LAC=0 のパズルが最も難しいと考えられる。このクラスのパズルは初期イデアル=戦術閉包という際立った特徴を持ち、今後、数学的な研究をしていきたい。

参考文献

- [1] Cannon, J., Bosma, W., Fieker, C., Steel, A.(eds.), HANDBOOK OF MAGMA FUNCTIONS Version 2.22, June 9, 201
- [2] 郷内邦義: 脳を鍛える数字パズルナンプレ (初級篇, 中級篇, 上級篇, 超上級篇, 難題, 超難問篇, 限界篇), 廣済堂出版, 2008.
- [3] Inkala, A.: AI Sudoku Top 10 set: 1, http://www.aisudoku.com/en/AIsudoku_Top10s1_en.pdf, 2006, 最終アクセス 2020 年 12 月 23 日.
- [4] Inoue, S.: Efficient Singleton set constraint solving by Boolean Groebner bases, Communications of JSSAC 1 (2012), 27-37.
- [5] Inoue, S. and Sato, Y.: A Mathematical Hierarchy of Sudoku Puzzles and Its Computation by Boolean Groebner Bases, LNCS 8884, Springer, 2014, 88-98.
- [6] 丸山陽太郎, 井上不変量を用いた数独型パズルの難易度判定, 東京電機大学大学院理工学研究科修士論文, 2019.
- [7] Nakano, T., and Tonegawa, Y.: Introduction to Boolean Groebner bases and their applications to puzzles of Sudoku type, Journal of Algebra and Applied Mathematics, 12, 2014, 1-31.
- [8] Nakano, T., Arai, K. and Watanabe, H.: On the Inoue invariants of the puzzles of Sudoku type, Communications of JSSAC 2 (2016), 1-14.
- [9] Nakano, T., Minami, S., Harike, S., Arai, K., Watanabe, H., and Tonegawa, Y.: On the Inoue invariants of the puzzles of sudoku type II, Bulletin of JSSAC 24 (2018), 77-90.
- [10] Nakano, T., Maruyama, Y., and Ohki, S.: On the Mathematical Evaluation of Difficulty Level of Sudoku Puzzles by Boolean Groebner Bases, Far East Journal of Applied Mathematics 106 (2020), 43-70.
- [11] Sato, Y., Inoue, S., Suzuki, A., Nabeshima, K., and Sakai, K.: Boolean Groebner bases, J. Symbolic Comput. 46 (2011), 622-632
- [12] Tarek.:<http://forum.enjoysudoku.com/the-hardest-sudokus-new-thread-t6539.html>
- [13] 落合啓之, 示野信一ほか, グレブナ基底と代数多様体入門 (上), 丸善出版, 2000, 4.
- [14] 大木茂樹, Boolean Groebner 基底を用いた数独型パズルの 2 つの数学的難易度判定指標について, 東京電機大学大学院理工学研究科修士論文, 2020.
- [15] 中野哲夫, 進藤未来, 御子崇尚希, 吉原元, 数独解を求める CII アルゴリズムの最近の発展と応用-不変量 SMYI と MDSL 予想-, 日本数式処理学会発表, 2022.6.18.
- [16] Tetsuo Nakano, Miku Shindou, Naoki Mikoshiba and Tsukasa Yoshihara, THE SMY INVARIANT AND THE MDSL CONJECTURE IN THE CII ALGORITHM FOR SOLVING SUDOKU PUZZLES, Far East Journal of Applied Mathematics 114(2022),25-48.