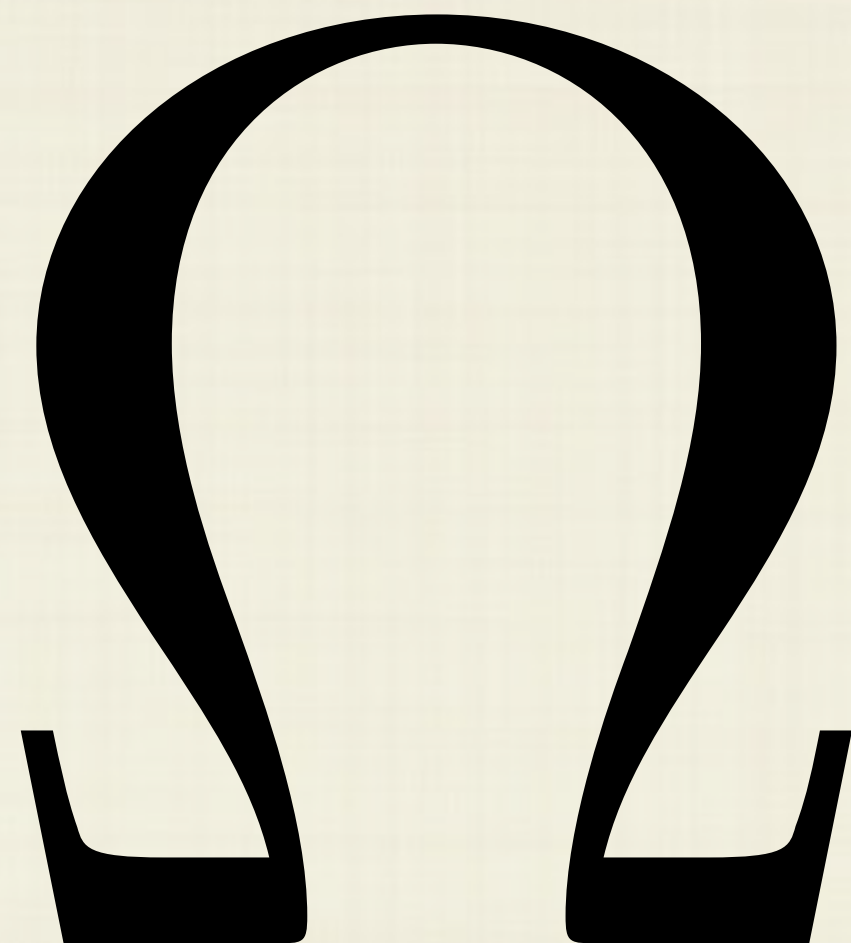


数学でランダム性と 向き合う

宮部賢志 @ 明治大学数学科

2017年10月7日(土) MATHPOWER



$$\Omega = \sum \{2^{-|\sigma|} : U(\sigma) \downarrow\}$$

ここで, U は接頭万能機械.

.

$\Omega \in (0, 1)$: 停止確率 (Halting Probability)

- 数学のすべての未解決問題の解答を知っている数
- Ω はランダムで、最初の有限桁を圧縮することはできない

Ω は知の数 (Number of Wisdom) とも呼ばれる.

■ 第1部： Ω は賢い

■ 第2部： Ω はランダム

■ 第3部：ランダム性と人工知能

第1部：Ωは賢い

第1部目次

- アルゴリズムとは何か
- 万能機械の存在
- 計算枚挙可能集合 (c.e.集合)

問題. $a, b \in \mathbb{N}$ として,

$$ax + by = 1$$

を満たす整数の組 (x, y) を求めるアルゴリズムを与えよ.

解答. 拡張ユークリッドの互除法で高速に求めることができる. □

問題. n を 3 以上の整数として,

$$x^n + y^n = z^n$$

を満たす自然数の組 (x, y, z) を求めるアルゴリズムを与えよ.

解答. フェルマーの最終定理により, そのような組は存在しない. □

問題 (ヒルベルトの第 10 問題). ディオファントス方程式 (整係数多変数高次不定方程式) を解くアルゴリズムを与えよ. (整数解が存在するかどうか判定し, 存在するならばその解を与える.)

MRDP 定理 (マチャセビッチ, ロビンソン, デイビス, プットナム, 1970) による解答.

解答. そのようなアルゴリズムは存在しない.

□

不可能性

定理（ゲーデル 1930）.

自然数論を含む帰納的公理化可能な理論は、
証明も反証もできない命題が存在する.

定理（チューリング 1936）.

プログラムが停止するかどうかを
判定するプログラムは存在しない.

同じ形の問題

- 可算個（自然数の個数）の候補
- 成立判定は有限時間で可能
- 存在するかどうかが問われている
- 半決定可能（semi-decidable） または
計算枚挙可能（c.e., computable enumerable）

ディオファントス方程式を解く困難さ

=

数学の主張が証明可能かどうかを判定する困難さ

=

プログラムが止まるかどうかを判定する困難さ

定義 (非形式的な定義). 関数 $f : \subseteq 2^{<\omega} \rightarrow 2^{<\omega}$ が部分計算可能であるとは, 適当なプログラミング言語でその関数を表現できることをいう.

そのような関数のことを機械と呼ぶこともある.

例. 以下の関数は計算可能.

- (i) $\sigma \in 2^{<\omega}$ に対し, $f(\sigma) = \sigma\sigma$
- (ii) $n \in \omega$ に対し, $f(n) = n^2$

自然数と 2 進有限列は同一視する.

入 0 1 00 01 10 11 000

1 2 3 4 5 6 7 8

自然数

2進数

文字列

6 $\longleftrightarrow$ 110 $\longleftrightarrow$ 10

最初の1と
とる

定理. 以下の性質を満たす **万能な部分計算可能関数** $U : \subseteq 2^{<\omega} \rightarrow 2^{<\omega}$ が存在する : 任意の部分計算可能関数 f に対し, ある文字列 $\sigma \in 2^{<\omega}$ が存在して, すべての $\tau \in 2^{<\omega}$ で,

$$U(\sigma\tau) \simeq f(\tau)$$

が成り立つ.

$\simeq$ は左辺が定義されているならば右辺も定義されており一致する. 左辺が定義されていないならば右辺は定義されていない.

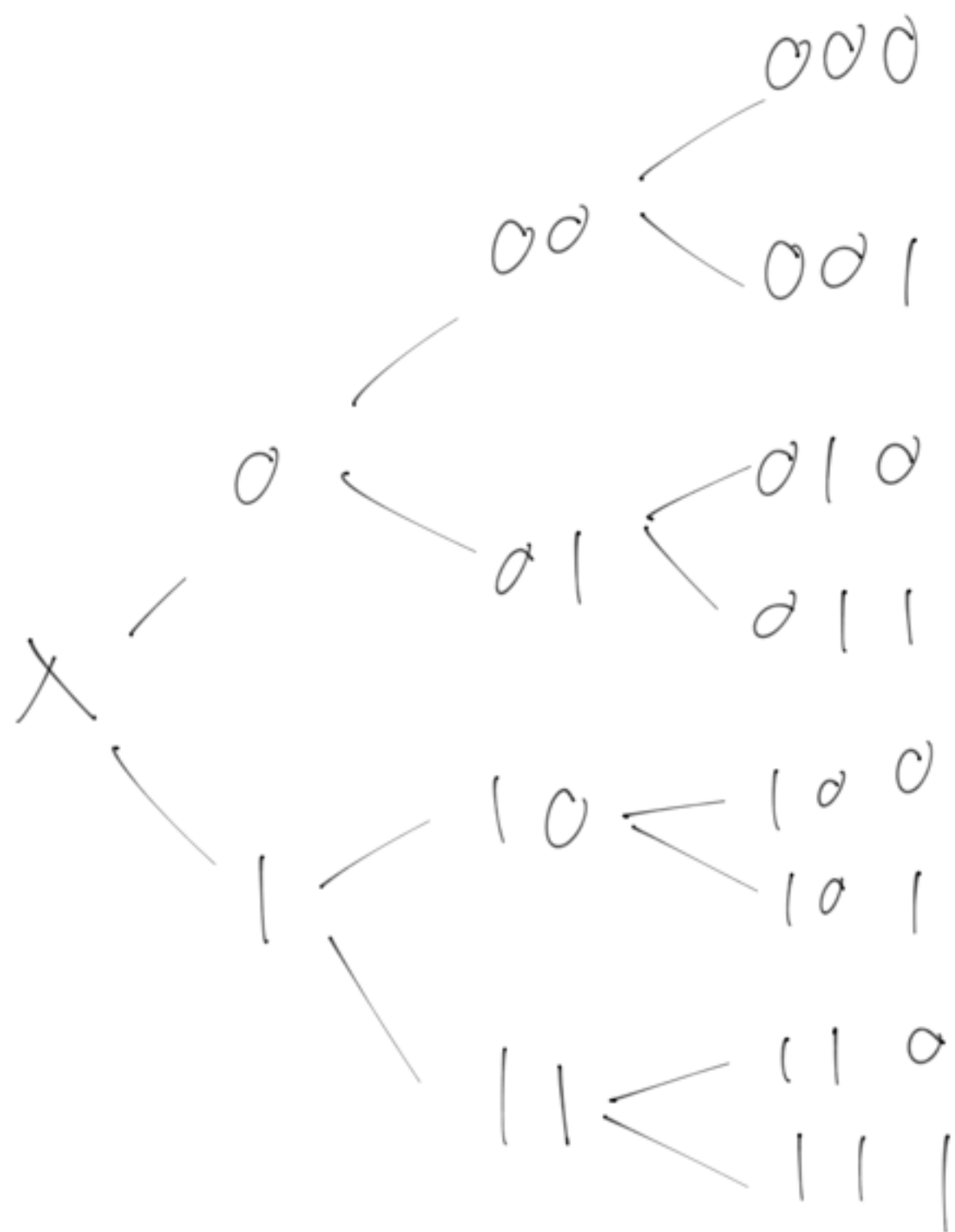
要するに

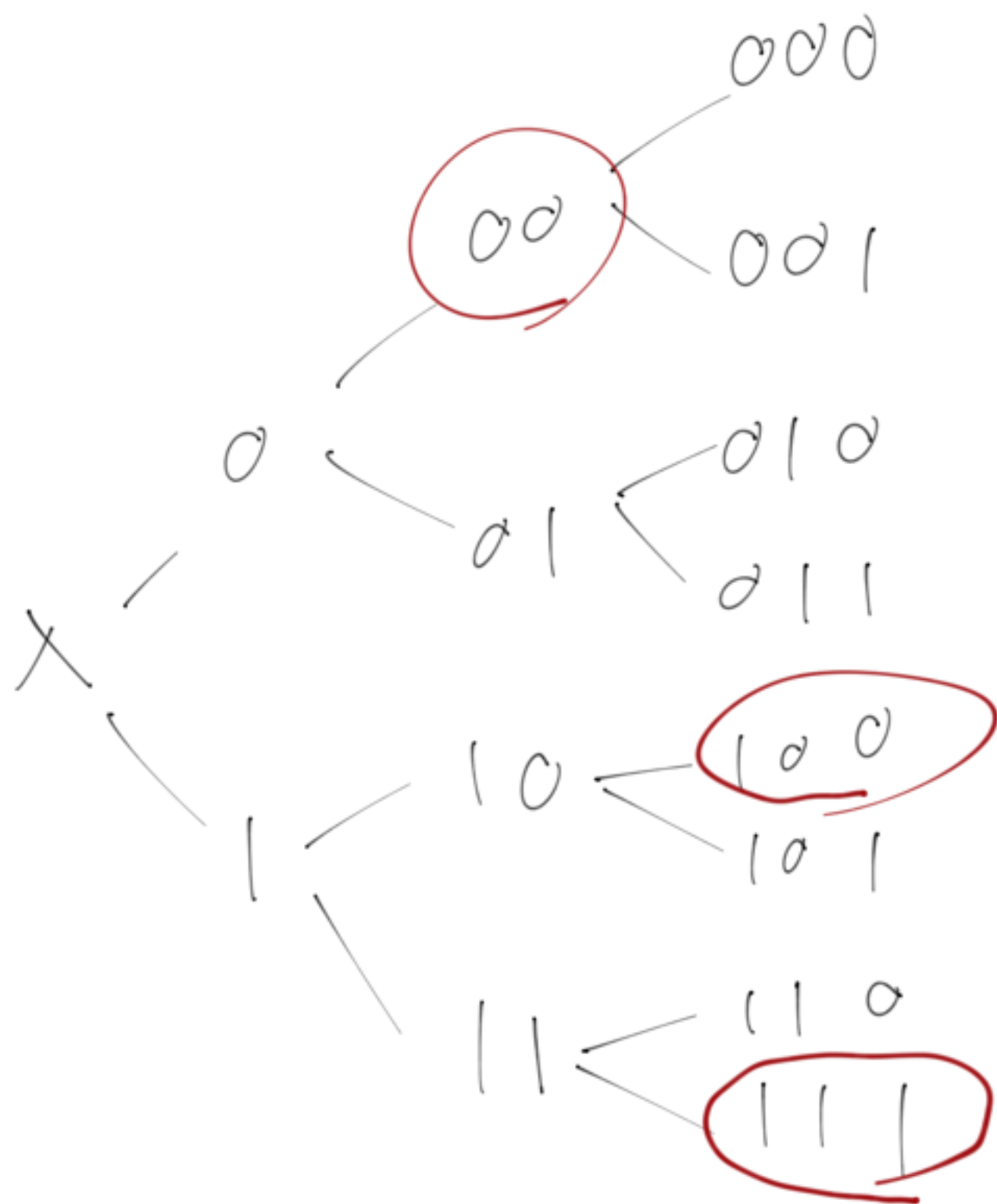
- $U = \text{PC}$, $\sigma = \text{コンパイラ}$, $\tau = \text{プログラム}$
- 入力文字列 = プログラム.
(但し出力は1つに決まる. 0および1での記述.)
- 万能 = 他の言語の模倣ができる.
特に止まらないプログラムも模倣できる.

定義. 機械 M が接頭機械 (prefix-free) とは, その定義域に, 接頭辞になっている異なる 2 つの文字列が存在しないことをいう:

$$\sigma \neq \tau \in \text{dom}(M) \Rightarrow \sigma \not\leq \tau$$

どこでプログラムが終わりなのか分かる. 普通は EOF を置く. それ以外の方法でも良い.

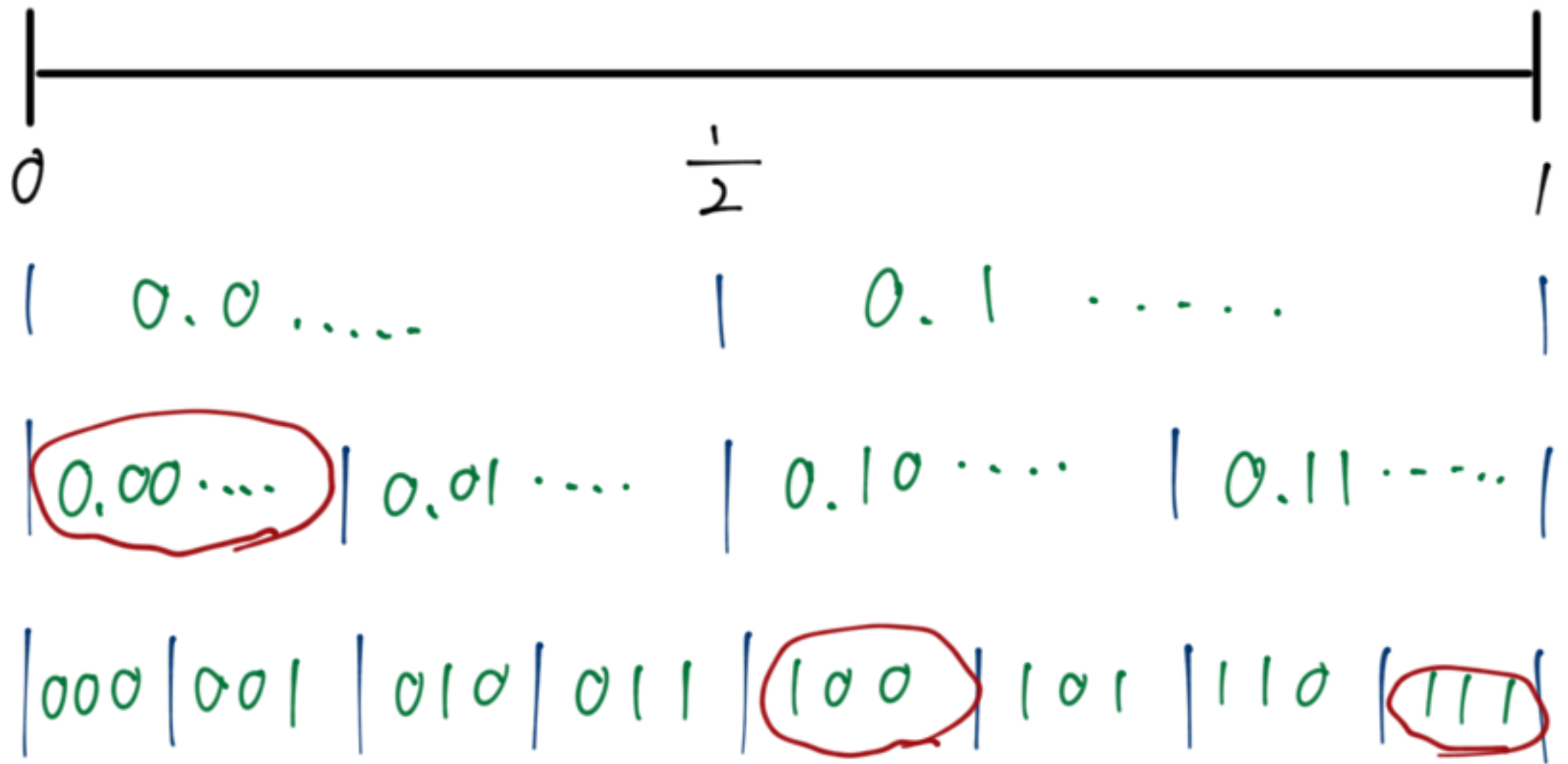




$$\Omega = \sum \{2^{-|\sigma|} : \sigma \in 2^{<\omega}, U(\sigma) \downarrow\}$$

ここで、 U は接頭万能機械.

接頭機械であることから、 $\Omega \in [0, 1]$ が出て来る (Kraft
の不等式).



Q

Q

Q

定義. 自然数の集合 A が **c.e.** であるとは, ある計算可能な関数 f が存在して,

$$A = \text{rng}(f)$$

となることをいう.

定義. 実数 α が**左 c.e.** であるとは,

$$\{q \in \mathbb{Q} : q < \alpha\}$$

が c.e. であることをいう.

定理. Ω は $H = \{\sigma : U(\sigma) \downarrow\}$ を計算できる.

証明. それぞれの $\sigma \in 2^{<\omega}$ に対して, $\Omega - \Omega_s < 2^{-|\sigma|}$ となる s を探す. $U(\sigma)[s] \downarrow$ であればもちろん $U(\sigma) \downarrow$ である. $U(\sigma)[s] \uparrow$ であれば, $U(\sigma) \uparrow$ である. もし, ある $t > s$ で $U(\sigma)[t] \downarrow$ であれば,

$$\Omega \geq \Omega_t \geq \Omega_s + 2^{-|\sigma|}$$

となり矛盾するからである.

□



Ω_s

$\updownarrow < 2^{-161}$



Ω

第1部まとめ

- 計算可能性理論 (Computability Theory) 面白い

第2部： Ω はランダム

ランダムとは？

- 規則がない，予測が不可能，無作為 (wikipedia)
- 確率変数，random variable (確率論)
- 無作為抽出 (統計学)

twitterより

- ランダムに発送する（どれかは分からない）
- ランダムに移動する（適当に，気ままに）
- ランダムに敵が出現する（間隔が一定でない？）

どれがランダム？

- 01...
- 00100100001111110110101010001000100001011010...
- 01100000001000011011010001101110101100100010...

どれがランダム？

- 01...
=>01の繰り返し
- 00100100001111110110101010001000100001011010...
=>円周率 π の小数部分の2進展開
- 01100000001000011011010001101110101100100010...
=>R言語で作った疑似乱数

ポイント

- 確率で考えるとどれも同じで区別できない
- 一見するとランダムでも、規則が分かればランダムでないように見える
- 短く表現できるかどうかでランダム性を測る

定義. 文字列 $\sigma \in 2^{<\omega}$ の Kolmogorov 複雑性を,

$$K(\sigma) = \min\{|\tau| : U(\tau) = \sigma\}$$

で定義する. ここで, $U : \subseteq 2^{<\omega} \rightarrow 2^{<\omega}$ は接頭万能機械である.

[illegible]

$$1\text{MB} = 8 * 10^6 \text{bit}$$

1	2	3	4	5	6	7	8	9	10	...
λ	0	1	00	01	10	11	000	001	010	...

という対応関係を取ると, $n \in \omega$ を表現するには, だいたい $\log_2 n$ くらいあれば良い. 接頭表現であっても大きくは変わらない.

定義. $X \in 2^\omega$ が **ランダム** であるとは,

$$K(X \upharpoonright n) > n - O(1)$$

であることを言う.

注意. この概念は Martin-Löf ランダムネスと呼ばれている. 1960s~70s の Martin-Löf, Kolmogorov, Levin, Schnorr, Chaitin らの研究による.

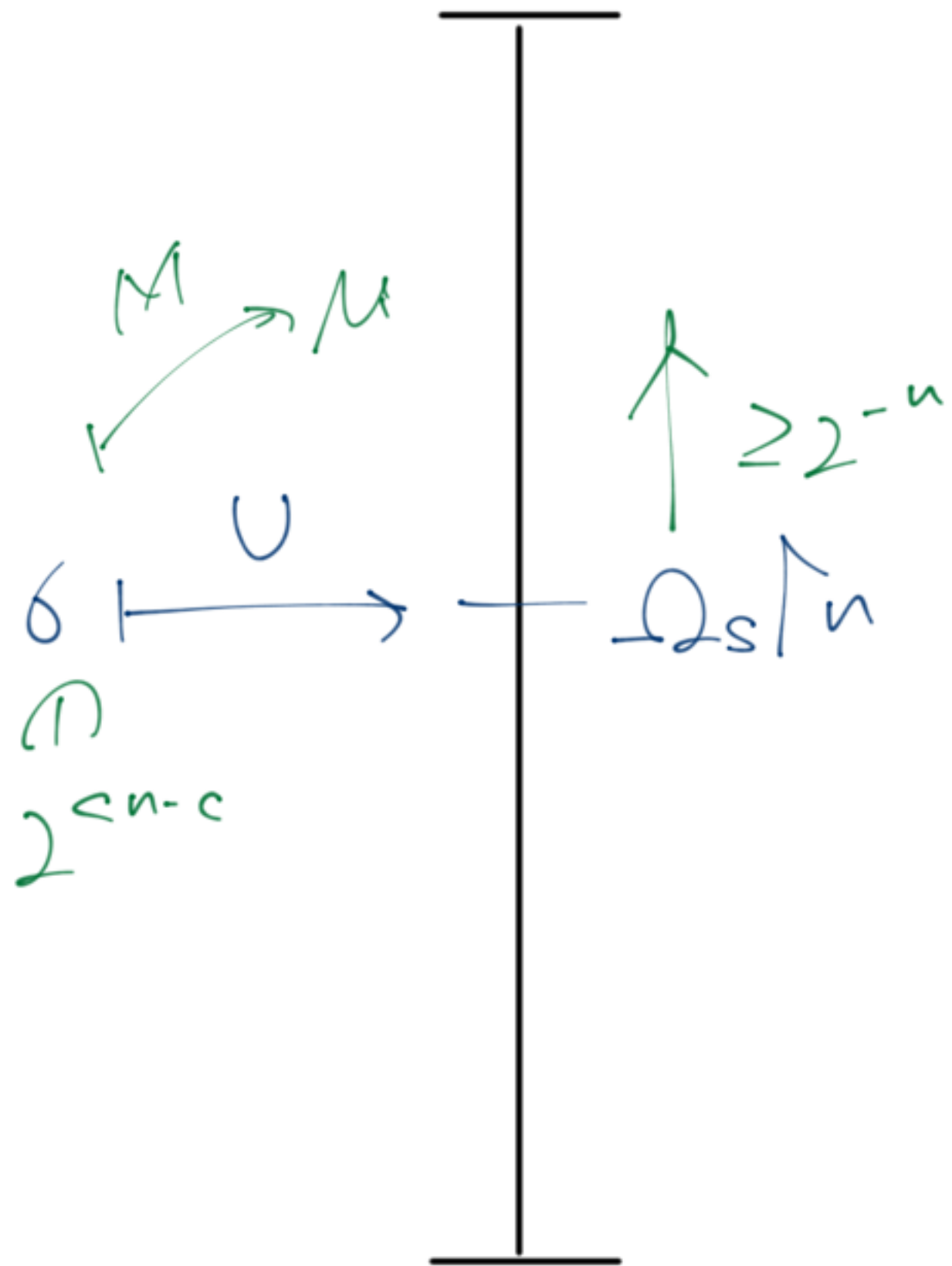
定理. Ω はランダムである.

証明. M のプログラムを p として, $\sigma \in 2^{<n-|p|}$ に対し $U(\sigma) = \Omega_s \upharpoonright n$ となる (σ, s) の組を探し, $M(\sigma) = \mu \notin \text{rng}U[s]$ を宣言する.

$U(p\sigma)[s] \uparrow$ だが $U(p\sigma) \downarrow$ なので,

$$\Omega - \Omega_s \geq 2^{-(|p|+|\sigma|)} \geq 2^{-n}$$

すなわち, $|\sigma| < n - c$ なら $U(\sigma) \neq \Omega \upharpoonright n$. □



ベリーのパラドックス

- 「19文字以内で記述できない最小の自然数」

$\Omega \in (0, 1)$: 停止確率 (Halting Probability)

- 数学のすべての未解決問題の解答を知っている数
- Ω はランダムで, 最初の有限桁を圧縮することはできない

Ω は知の数 (Number of Wisdom) とも呼ばれる.

補足

定理. $\Omega \equiv_T \emptyset'$ である. しかし, Ω は $\leq_{tt}$ に関しては完全ではない. 特に $\emptyset' \leq_{tt} A$ であれば, A は *Kurtz* ランダムではない.

定理. U を万能機械とする. Ω_U は左 *c.e.* で *ML* ランダムである. 逆に, 任意の左 *c.e.* な *ML* ランダム α に対し, ある万能機械が存在して $\alpha = \Omega_U$ となる.

第2部まとめ

- アルゴリズム的ランダムネスの理論
(Algorithmic Randomness) 面白い！

ランダム性と 人工知能

ルール

- 次のビットが0か1かに所持金の一部分を賭ける
- 当たれば2倍になって返ってくる
はずれれば没収される
- 借金ができないものとする

ランダム=予測不可能

- 上記のルールของเกมにおいて,
どんな左c.e.の（下側から計算可能な）戦略でも,
所持金を無限に大きくできない列
- MLランダムな列

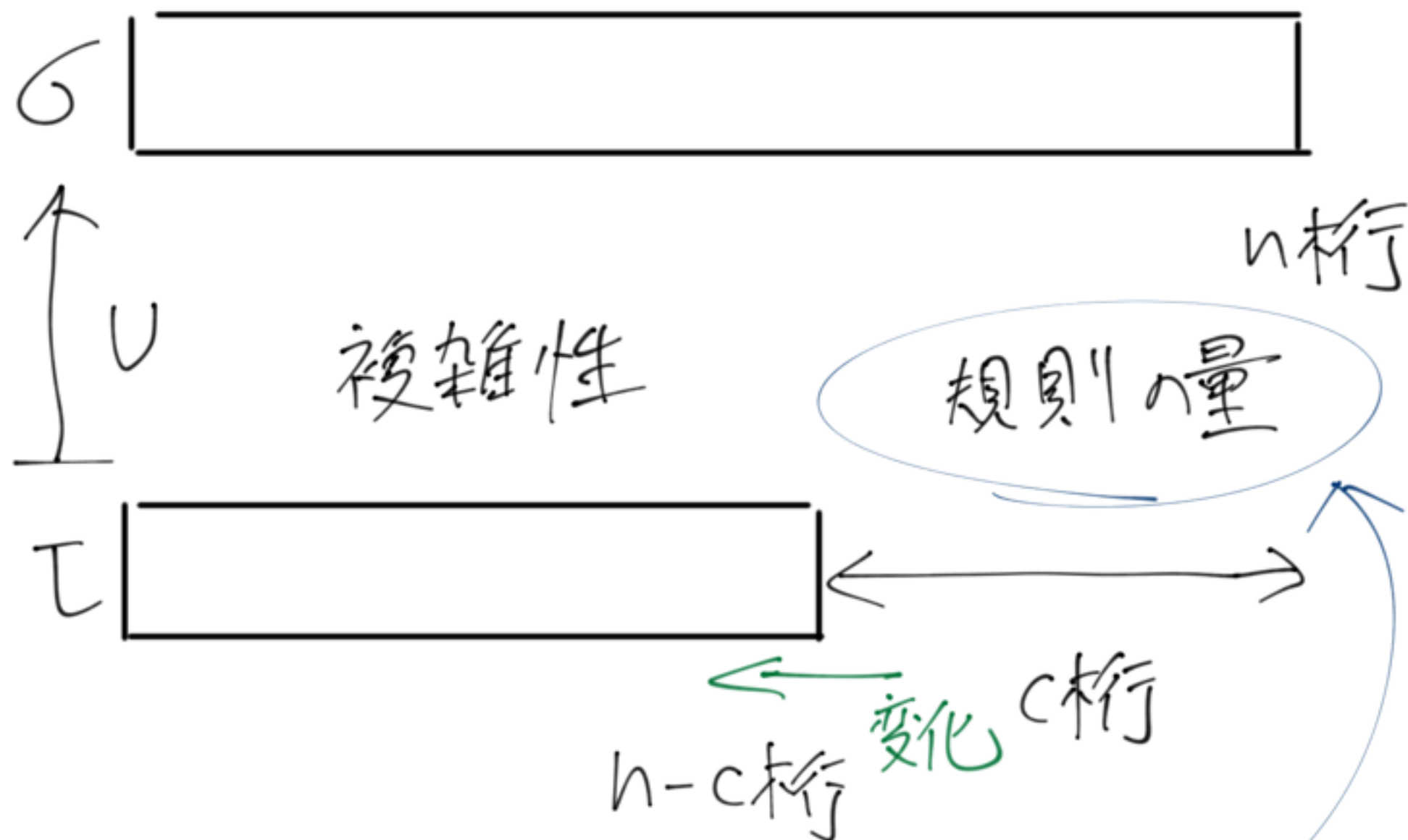
非ランダム性＝予測可能性

機械予測

=

機械に非ランダム性を見つけさせる

予測と理解の関係は？

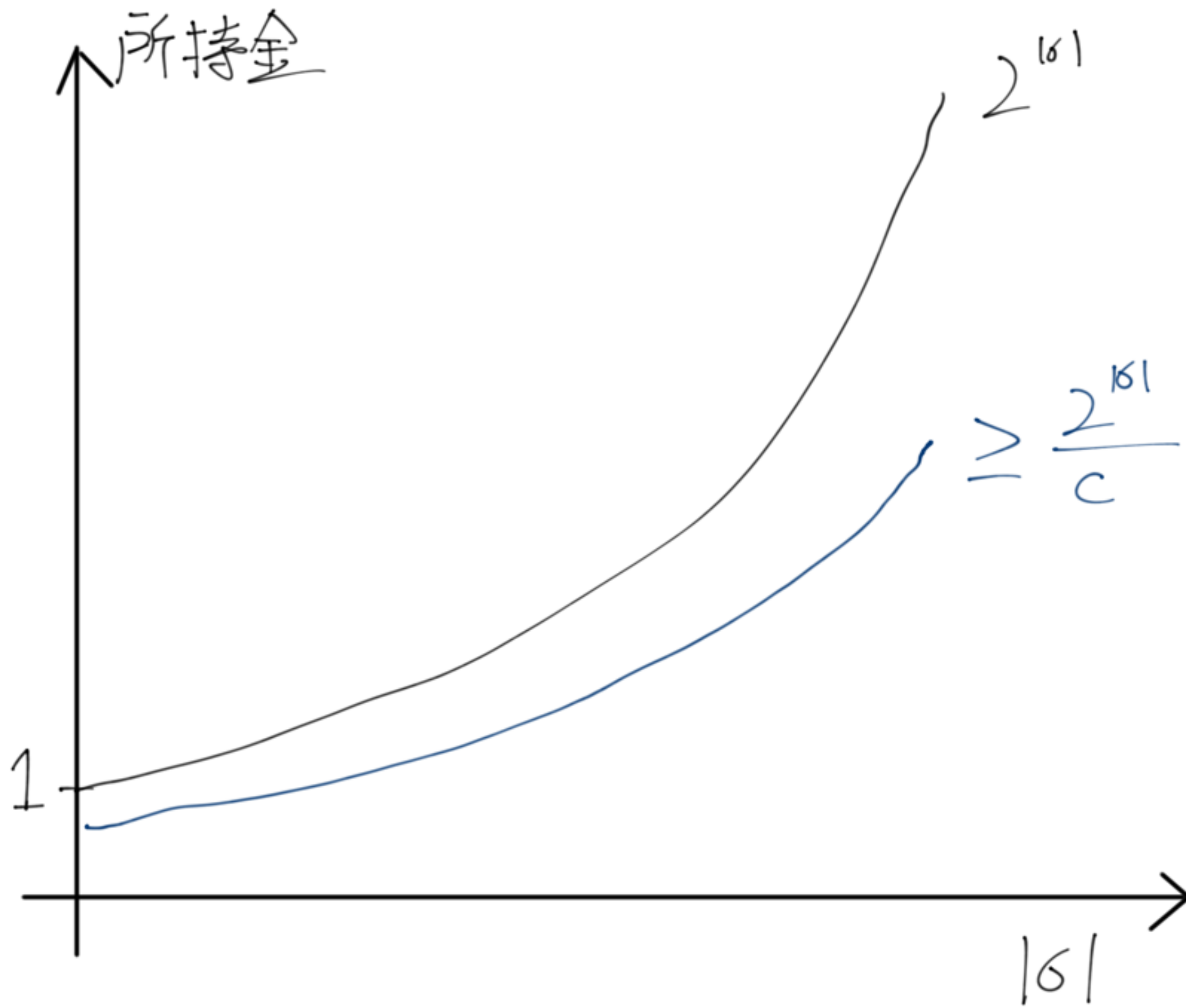


randomness deficiency
ランダム性不足量

Solomonoff の万能推論 :

$$M(\sigma) = \sum \{2^{-|\tau|} : U(\tau) = \sigma*\}$$

すべての左 c.e. の戦略と比べて, 定数倍を除いて悪くない.



補足

定理. 任意の計算可能な列 $x \in 2^\omega$ に対し,

$$1 - M(x_n | x_{<n}) = O(1) \times 2^{-K(n)}$$

が成立する.

$n - 1$ 桁まで知った時に, n 桁目で正しい値 x_n へ賭ける割合は, n が大きくなると 1 に近づく. その割合と 1 との差はだいたい $2^{-K(n)}$ くらい.

ほとんどの複雑な n については $\frac{1}{n}$ より小さくなるが, 単純な n では $\frac{1}{n}$ よりもかなり大きくなる.

第3部まとめ

- アルゴリズム的確率 (Algorithmic Probability)
面白い

まとめ

- 計算可能性という観点から、
確率や予測の概念を眺めてみると、
いろいろと面白いことが分かる。