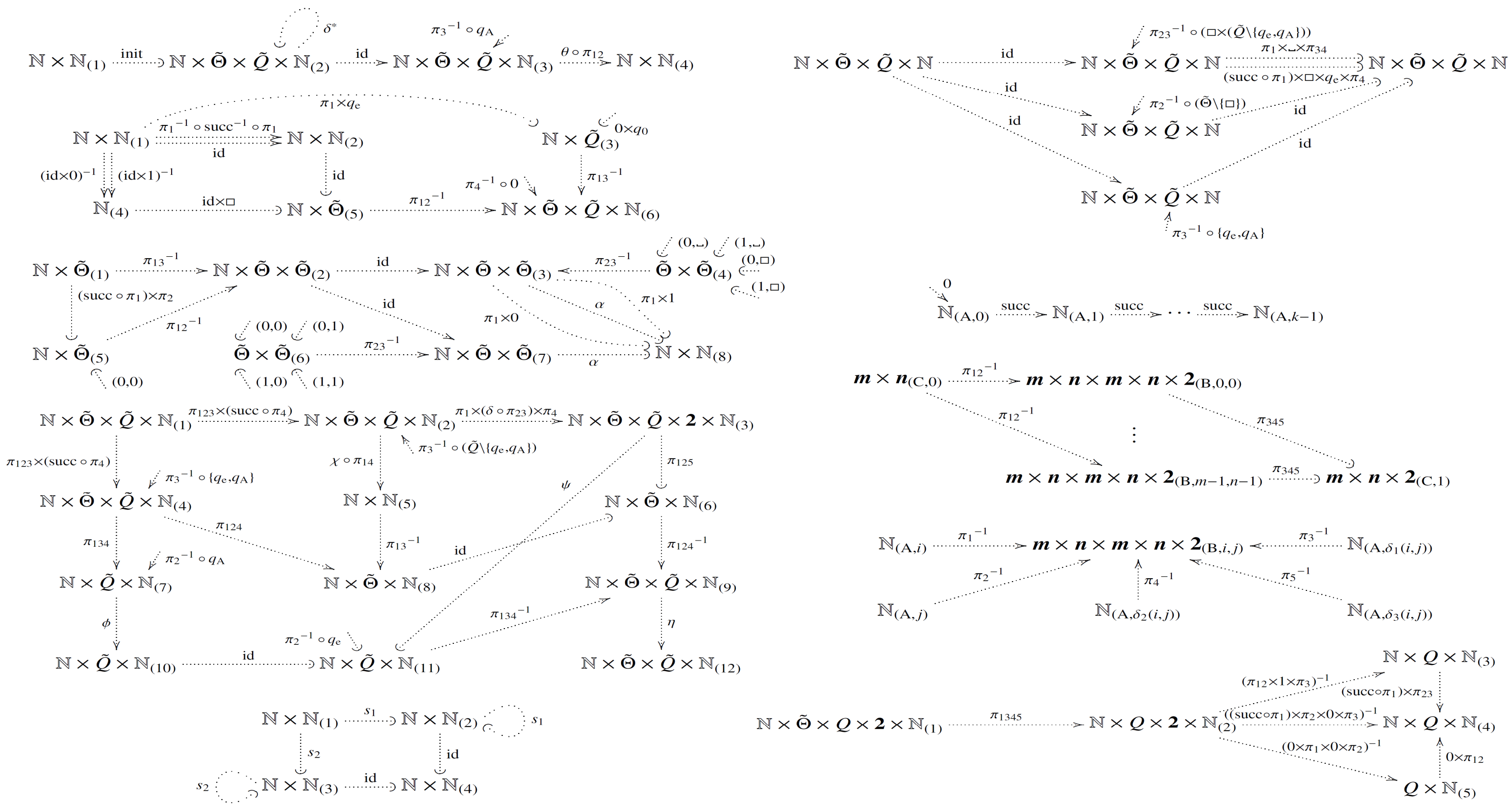
パターンとは何か
非記号計算と一般対象の情報計量

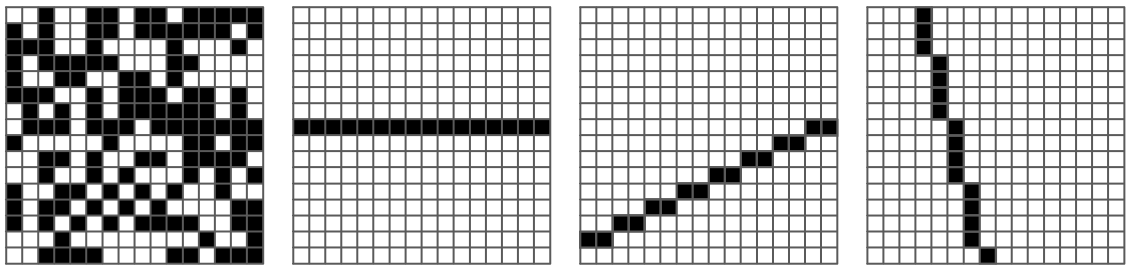
石川 博

名古屋市立大学
大学院システム自然科学研究科



高次元データとパターン発見

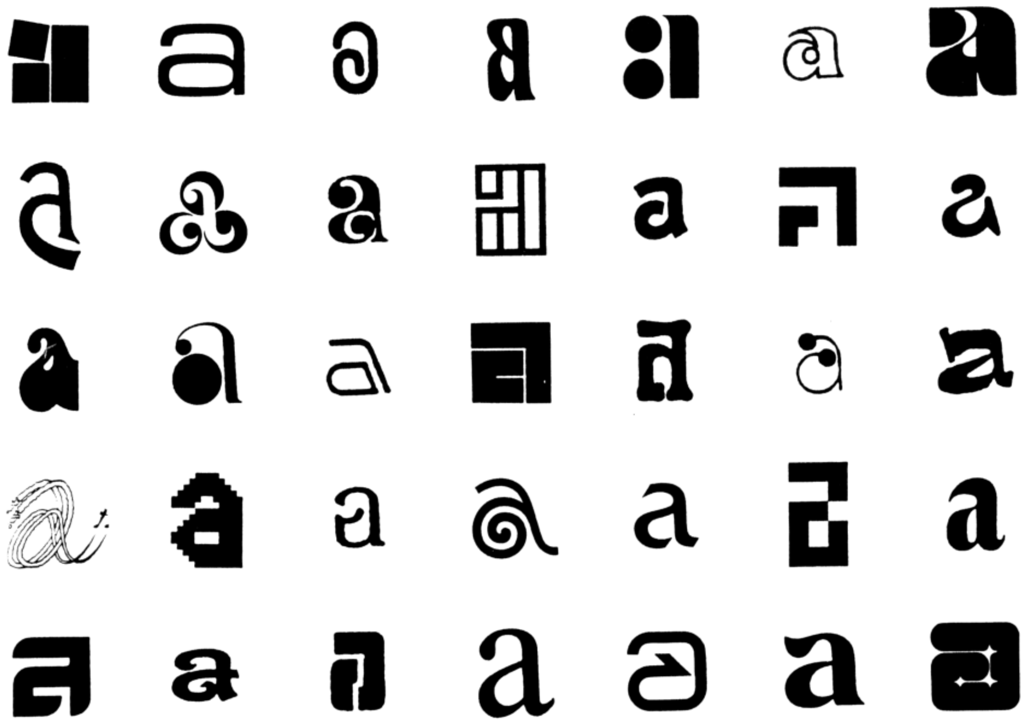
- 高次元データ
  - 個々の数値は意味を持たない
  - 値の特定の組み合わせに情報がある

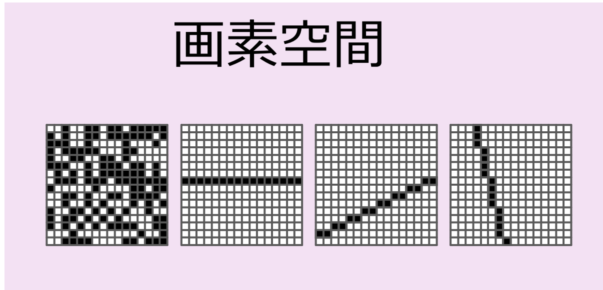


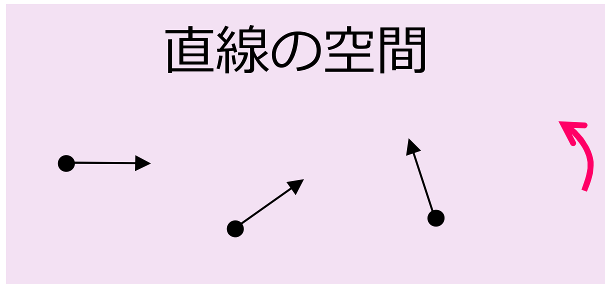
- 探すもの（パターン）がわかっている必要
- 古典的パターン認識
  - パターンはプログラムで表現
  - 人に「見える」ものに限定される

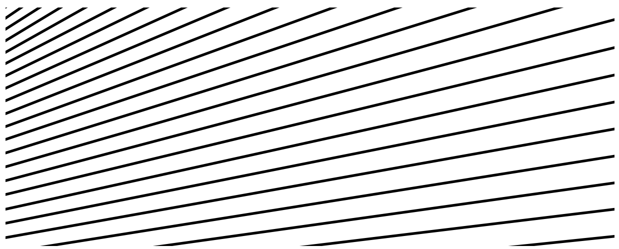
パターンの階層

- パターンは生データレベルとは限らない
  - 階層的な表現・認識 → パターンの空間









Minimum Description Length

- 「オッカムの剃刀」の形式化
- 「世界」のモデルを学習するために、
  2つの部分に分けてそれを記述する:
  1. 対象に規則性があればアルゴリズムでモデル化し圧縮する
  2. 残りの非規則性はデータとして持つ
- 2部分を合わせて最も短くなるのが最良のモデルである。

記号列の場合

ある文字列を別の文字列で表す2方法:

A: そのまま

B: その文字列を書き出すプログラムで

Bは文字列中のパターンを利用できる

01010101010101010101010101010101…

01011011101111011111011111110…

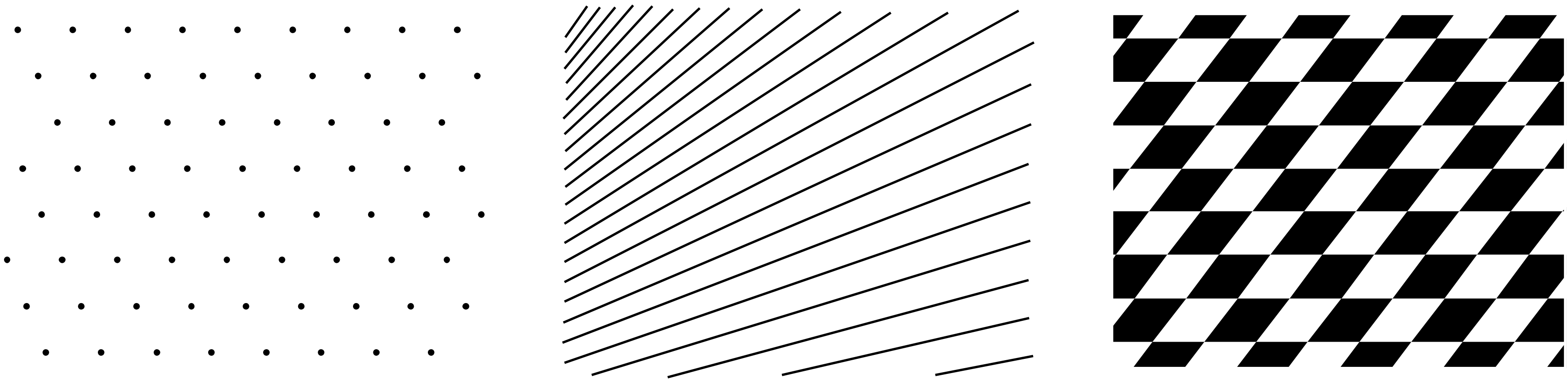
パターンがある　⇔　Bで短く表現

Bで最短の長さ→　コルモゴロフ複雑性

パターンとは計算(アルゴリズム)である

記号列以外の場合

これらの「パターン」の存在をどのように
定義したらよいだろうか。

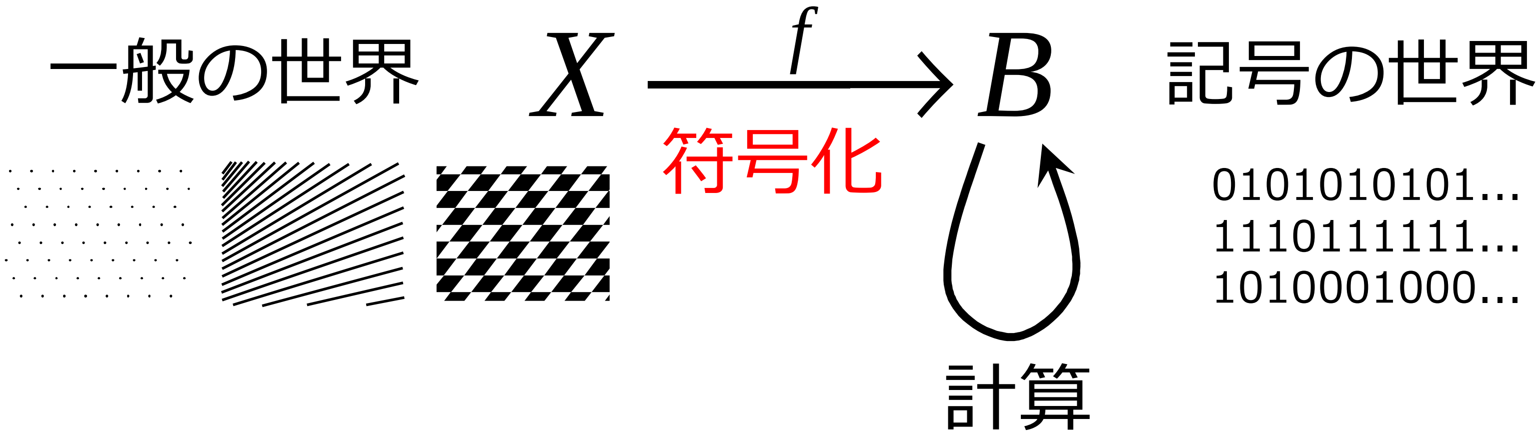


- コルモゴロフ複雑性のような、何らかの
情報計量を定義

計算機科学のテーゼ

“情報は符号化できる”

- 計算は記号の世界(＝ビット列の世界)でしか定義されていない
- その他の世界のことについて計算をするときには、まずその世界を符号化(encode)する
- 符号化は任意



記号の世界とその他の世界

一般の対象の情報計量を定義するには

- 符号化を固定する
- 下のように一般の対象のコルモゴロフ複雑性を記号列のそれ(K)を使って定義する

一般の世界

$$\begin{array}{ccccc} X & \xrightarrow[\textbf{符号化}]{f} & B & \xrightarrow{K} & \mathbb{N} \\ \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\ x & & f(x) & & K(f(x)) \end{array}$$

符号化の限界

- 符号化は本当に任意ではない
- 符号化が任意なら情報計量も任意
- 実際に符号化を定義する必要

任意の置換 p

一般の世界

$$\begin{tikzcd}
 & & K(p\circ f(x)) \arrow[d, phantom, sloped, "\in"] \\
X \arrow[r, "f"] & B \arrow[loop above, "p"] \arrow[r, "K"] & \mathbb{N} \\
x \arrow[u, phantom, sloped, "\in"] & f(x) \arrow[u, phantom, sloped, "\in"] & K(f(x)) \arrow[u, phantom, sloped, "\in"]
\end{tikzcd}$$

符号化

符号化の限界

符号化は

- 自然であることが必要
  - –対象の規則性が符号の規則性に対応
- 例：ユークリッド空間の点

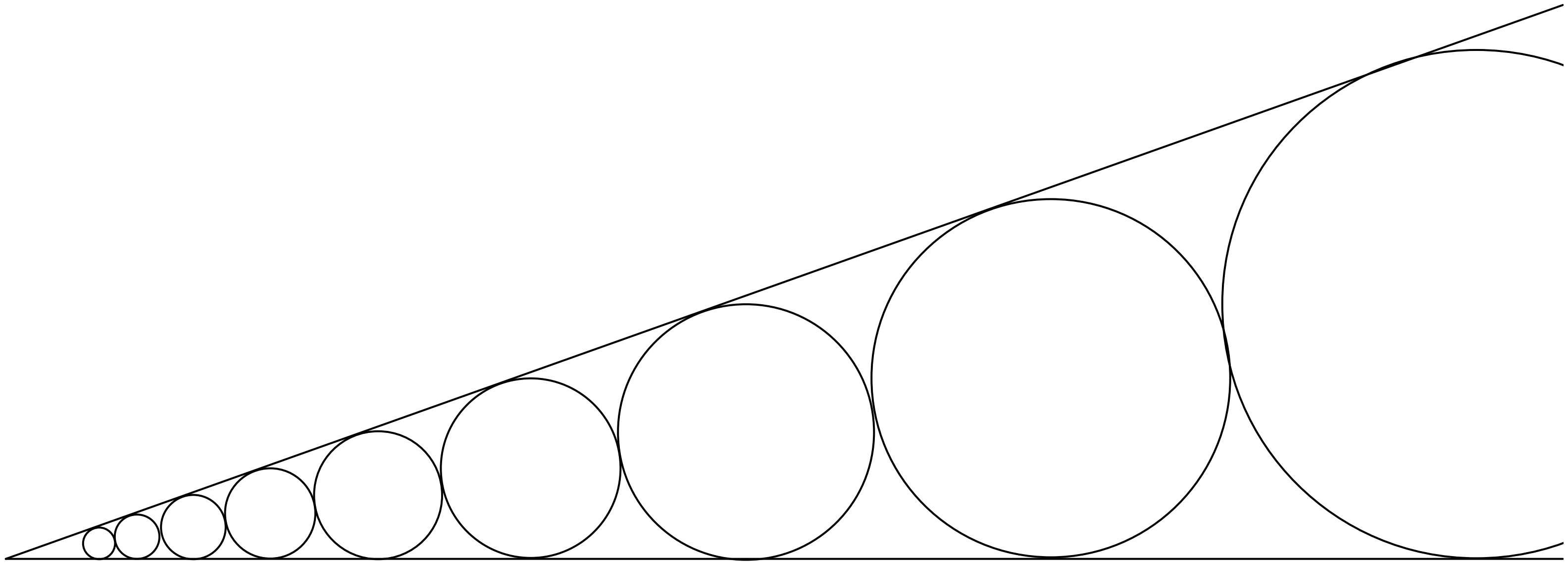
  点 → 座標 → 文字列
  - –1点でも無限の情報量
  - –無限個の点でも有限の情報量

    (3, 0), (14, 0), (159, 0), (2653, 0), …
  - –符号化が変われば意味のなくなる規則性

符号化の限界

符号化は

- 自然であることが必要
  - 対象の規則性が符号の規則性に対応
  - 対象同士の関係も符号化が必要

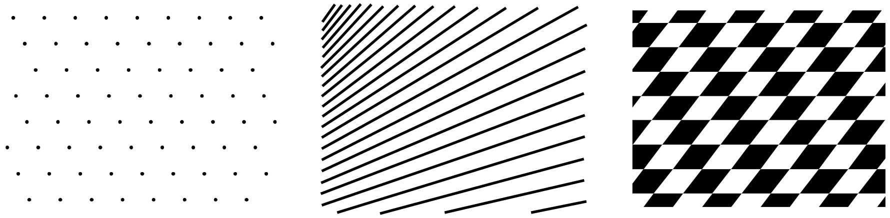


符号化の限界

符号化は

- 自然であることが必要
  - 対象の規則性が符号の規則性に対応
  - 対象同士の関係も符号化が必要
  - 符号化～座標系
- 有限であることが必要
  - 有限ならば符号化間の翻訳は無視できる
  - 無限だと符号化自体の持つ情報量が問題

例：有限情報を持つ図形



符号化

- 実数：2進展開
- 点: 実数2つ
- 直線: 点２つ
- 円: 点と実数
- ……

例：有限情報を持つ図形

しかし:

- どこまで定義するのか?
  - 全ての可能な図形を尽くすことはできない
- 何が「有限情報を持つ図形」であるか先に決めている
  - 「ここに列挙したもの」というだけ
- 点の集合だけを符号化して後は計算で定義
  - 文字列が足りない！

符号化アプローチの問題点

- 符号化自体をアルゴリズムで定義したい
  → 符号化後にしか使えない

- 自然な符号化にしたい
  – 部分の情報量をアプリオリに指定したい
  例：点の情報量は一定にしたい

$$\begin{array}{ccccc} X & \xrightarrow[\text{符号化}]{f} & B & \xrightarrow{K} & \mathbb{N} \\ \cup\!\!\!\mid & & \cup\!\!\!\mid & & \cup\!\!\!\mid \\ x & & f(x) & & K(f(x)) \end{array}$$

アプローチ

- 符号化自体をアルゴリズムで定義したい
  → 符号化後にしか使えない

  → 計算を符号化することなく直接定義

- 自然な符号化にしたい
  – 部分の情報量をアプリオリに指定したい
    例：点の情報量は一定にしたい

→ 基本構造の情報量をアプリオリに指定

アプローチ

- ある文字列を別の文字列で表す2方法:
  - そのまま (密)
  - その文字列を書き出すプログラムで(疎)

- 同様に、対象を２つの方法で表す:
  - 部分集合として (密)

    基底表現
  - 基底表現を「書き出す」新表現 (疎)

    図式 と 断面

基底表現

- 生のデータを表すビットマップのようなもの
- 抽象化 $\to$ 部分集合
  - 直線、円、$\cdots$ $\to$ 空間の部分集合
  - 画像 $\to$ $\mathbb{R}^2 \times$ *Color* の部分集合

    $D \subset \mathbb{R}^2$上の関数 $I(p) \to \{(p, I(p)) \mid p \in D\}$
  - その他　例：物体の表現

    空間を$\mathbb{R}^3$、各点を占める物質の集合をM

自転車 $\subset \mathbb{R}^3 \times M$, M={鉄、ゴム、ガラス、…}

IBIS2009

基底表現

- 生データ・信号レベル
- 近似は後で考えることにして、近似の対象をまず定義したい
- アプリオリ
- 構造は反映しない
- 表現可能なもののほとんどがノイズ

定義：図式と断面

- 集合とそれらの間の冪写像を指定(図式)

例

$$\varphi : \mathcal{P}(Y) \to \mathcal{P}(X)$$
$$\psi : \mathcal{P}(Z) \to \mathcal{P}(Y)$$
$$\eta : \mathcal{P}(W) \to \mathcal{P}(Y)$$
$$\kappa : \mathcal{P}(X) \to \mathcal{P}(W)$$

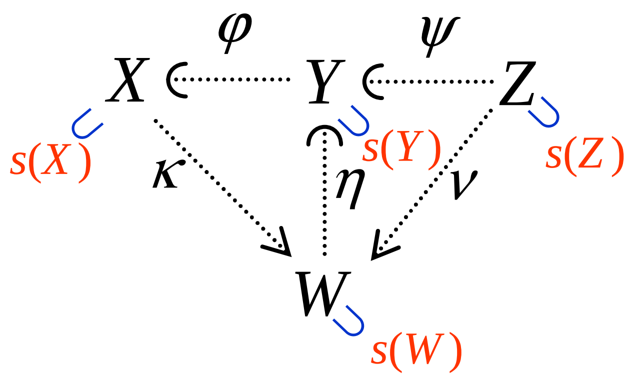
$$\begin{tikzcd}
X \arrow[dr, dotted, "\kappa"'] & Y \arrow[l, hook', dotted, no head, "\varphi"'] \arrow[d, hook, dotted, no head, "\eta"] & Z \arrow[l, hook', dotted, no head, "\psi"'] \arrow[dl, dotted, "\nu"] \\
 & W &
\end{tikzcd}$$

定義：図式と断面

- 集合とそれらの間の冪写像を指定(図式)

例

$$\varphi : \mathcal{P}(Y) \to \mathcal{P}(X)$$
$$\psi : \mathcal{P}(Z) \to \mathcal{P}(Y)$$
$$\eta : \mathcal{P}(W) \to \mathcal{P}(Y)$$
$$\kappa : \mathcal{P}(X) \to \mathcal{P}(W)$$



- 各集合にその部分集合を与える s を考える(断面)

定義：図式と断面

- 集合とそれらの間の冪写像を指定(図式)

例

$$\varphi : \mathcal{P}(Y) \to \mathcal{P}(X)$$
$$\psi : \mathcal{P}(Z) \to \mathcal{P}(Y)$$
$$\eta : \mathcal{P}(W) \to \mathcal{P}(Y)$$
$$\kappa : \mathcal{P}(X) \to \mathcal{P}(W)$$

$$\begin{tikzcd}
X \arrow[d, phantom, sloped, "\supset"] \arrow[drr, dotted, "\kappa"] & & Y \arrow[ll, hook', dotted, "\varphi"'] \arrow[d, phantom, sloped, "\supset"] & & Z \arrow[ll, hook', dotted, "\psi"'] \arrow[d, phantom, sloped, "\supset"] \arrow[dll, dotted, "\nu"] \\
s(X) & & W \arrow[u, hook, dotted, "\eta"'] \arrow[d, phantom, sloped, "\supset"] & & s(Z) \\
& & s(W) & &
\end{tikzcd}$$

$s(Y) \subset Y$

- 各集合にその部分集合を与える s を考える(断面)

- s は次を満たすこととする

$$s(S) = \bigcap_{\mu \in \mathrm{in}(S)} \mu(s(\mathrm{dm}(\mu)))$$
$$s(S) = \bigcup_{\mu \in \mathrm{in}(S)} \mu(s(\mathrm{dm}(\mu)))$$

$$s(X) = \varphi(s(Y))$$
$$s(Y) = \psi(s(Z)) \cup \eta(s(W))$$
$$s(W) = \kappa(s(X)) \cap \nu(s(Z))$$

$\mathrm{in}(S)$ ：集合S に入射する図式中の冪写像の集合

μ ： $\mathcal{P}(\mathrm{dm}(\mu)) \to \mathcal{P}(\mathrm{cdm}(\mu))$

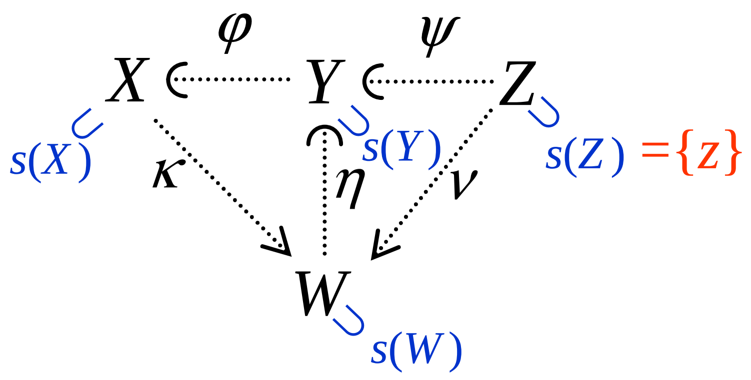
IBIS2009

定義：図式と断面

- 集合とそれらの間の冪写像を指定(図式)

例

$$\varphi : \mathcal{P}(Y) \to \mathcal{P}(X)$$
$$\psi : \mathcal{P}(Z) \to \mathcal{P}(Y)$$
$$\eta : \mathcal{P}(W) \to \mathcal{P}(Y)$$
$$\kappa : \mathcal{P}(X) \to \mathcal{P}(W)$$



- 各集合にその部分集合を与える s を考える(断面)

- s は次を満たすこととする

$$s(S) = \bigcap_{\mu \in \mathrm{in}(S)} \mu(s(\mathrm{dm}(\mu)))$$
$$s(S) = \bigcup_{\mu \in \mathrm{in}(S)} \mu(s(\mathrm{dm}(\mu)))$$

$$s(X) = \varphi(s(Y))$$
$$s(Y) = \psi(s(Z)) \cup \eta(s(W))$$
$$s(W) = \kappa(s(X)) \cap \nu(s(Z))$$

- いくつかの $s(\ )$ を指定する (部分断面)

定義：図式と断面

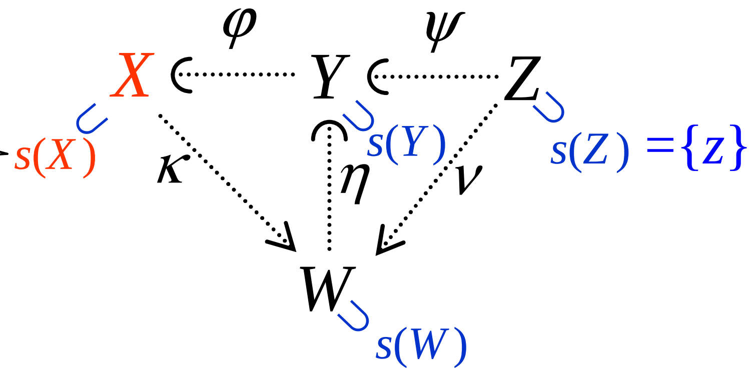
- 集合とそれらの間の冪写像を指定(図式)

例

$\varphi : \mathcal{P}(Y) \to \mathcal{P}(X)$

ここに対象が基底表現で与えられる

$\kappa : \mathcal{P}(X) \to \mathcal{P}(W)$



- 各集合にその部分集合を与える s を考える(断面)

- s は次を満たすこととする

$$s(S) = \bigcap_{\mu \in \mathrm{in}(S)} \mu(s(\mathrm{dm}(\mu)))$$

$$s(S) = \bigcup_{\mu \in \mathrm{in}(S)} \mu(s(\mathrm{dm}(\mu)))$$

$$s(X) = \varphi(s(Y))$$

$$s(Y) = \psi(s(Z)) \cup \eta(s(W))$$

$$s(W) = \kappa(s(X)) \cap \nu(s(Z))$$

- いくつかの $s(\ )$ を指定する (部分断面)
- 図式内の集合を一つ指定する

最も簡単な例

$$
\begin{tikzcd}
\{r\} \subset \mathbb{R} \arrow[r, dotted, "\mathrm{dist}^{-1}"] & X\times X' \supset s(X\times X') \arrow[d, dotted, "\pi_1"] \\
\{p\} \subset X' \arrow[ur, dotted, "\pi_2^{-1}"] & X \supset s(X)=?
\end{tikzcd}
$$

X, X' : ユークリッド平面

dist: $X\times X \to \mathbb{R}$:距離関数

π_1, π_2 : 射影

幕写像の記法

$f: A \to B$

$\Rightarrow$

$f: \mathcal{P}(A) \to \mathcal{P}(B) \qquad A \supset S \mapsto \{ f(x) \mid x \in S\} \subset B$

$f^{-1}: \mathcal{P}(B) \to \mathcal{P}(A) \qquad B \supset S \mapsto \{x \in A \mid f(x) \in S\} \subset A$

最も簡単な例

$$\begin{tikzcd}
\{r\} \arrow[r, phantom, "\subset"] & \mathbb{R} \arrow[r, dashed, "\mathrm{dist}^{-1}"] & X\times X' \arrow[d, dashed, "\pi_1"'] & s(X\times X') \arrow[l, phantom, "\subset"] \\
\{p\} \arrow[r, phantom, "\subset"] & X' \arrow[ur, dashed, "\pi_2^{-1}"] & X & s(X)=? \arrow[l, phantom, "\subset"]
\end{tikzcd}$$

X, X' : ユークリッド平面

dist: $X\times X \to \mathbb{R}$:距離関数

π_1, π_2 : 射影

$$s(X\times X') = \mathrm{dist}^{-1}(s(\mathbb{R})) \cap \pi_2^{-1}(s(X'))$$

$$s(X) = \pi_1(s(X\times X'))$$

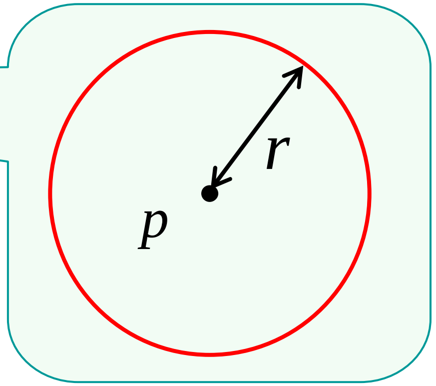
最も簡単な例

$$\begin{tikzcd}
 & \mathbb{R} \arrow[r, dotted, "\mathrm{dist}^{-1}"] & X\times X' \arrow[d, dotted, "\pi_1"] \arrow[r, phantom, "\supset"] & s(X\times X') \\
\{r\} \arrow[ur, phantom, "\subset"] & X' \arrow[ur, dotted, "\pi_2^{-1}"] & X & \\
\{p\} \arrow[ur, phantom, "\subset"] & & s(X)=? \arrow[u, phantom, "\subset"] &
\end{tikzcd}$$

X, X' : ユークリッド平面

dist: $X\times X\rightarrow\mathbb{R}$:距離関数

π_1, π_2 : 射影



$$s(X\times X') = \mathrm{dist}^{-1}(s(\mathbb{R}))\cap\pi_2^{-1}(s(X'))$$
$$= \{(x,y)\,|\,\mathrm{dist}(x,y)\in s(\mathbb{R}), \pi_2(x,y)\in s(X')\}$$
$$= \{(x,p)\,|\,\mathrm{dist}(x,p)=r\}$$

$$s(X) = \pi_1(s(X\times X'))$$
$$= \{x\,|\,\mathrm{dist}(x,p)=r\}$$

最も簡単な例

$$
\begin{tikzcd}
 & V \arrow[r, dotted, "\pi_1^{-1}"] & V\times\mathbb{R} \arrow[r, dotted, "\mathrm{mult}"] & V' \arrow[dl, dotted, "\mathrm{sub}^{-1}"'] & \\
\{v\} \arrow[ur, phantom, "\subset" sloped] & & & & \\
 & X' \arrow[r, dotted, "\pi_2^{-1}"] & X\times X' \arrow[r, dotted, "\pi_1"] & X & \\
\{p\} \arrow[ur, phantom, "\subset" sloped] & & & & ? \arrow[ul, phantom, "\supset" sloped]
\end{tikzcd}
$$

X, X' :ユークリッド平面

V, V' : 2次元ベクトル空間

mult: $V\times\mathbb{R} \longrightarrow V \quad (v, c) \mapsto cv$

sub: $X\times X \rightarrow V \quad (x, y) \mapsto x - y$

π_1, π_2: 射影

最も簡単な例

$$
\begin{tikzcd}
 & V \arrow[r, dotted, "\pi_1^{-1}"] & V\times\mathbb{R} \arrow[r, dotted, "\mathrm{mult}"] & V' \arrow[dl, dotted, "\mathrm{sub}^{-1}"'] \\
\{v\} \arrow[ur, phantom, "\subset" sloped] & & & \\
 & X' \arrow[r, dotted, "\pi_2^{-1}"] & X\times X' \arrow[r, dotted, "\pi_1"] & X \\
\{p\} \arrow[ur, phantom, "\subset" sloped] & & & ? \arrow[u, phantom, "\subset" sloped]
\end{tikzcd}
$$

X, X' :ユークリッド平面

V, V' : 2次元ベクトル空間

mult: $V\times\mathbb{R} \longrightarrow V \quad (v,c) \mapsto cv$

sub: $X\times X \longrightarrow V \quad (x,y) \mapsto x-y$

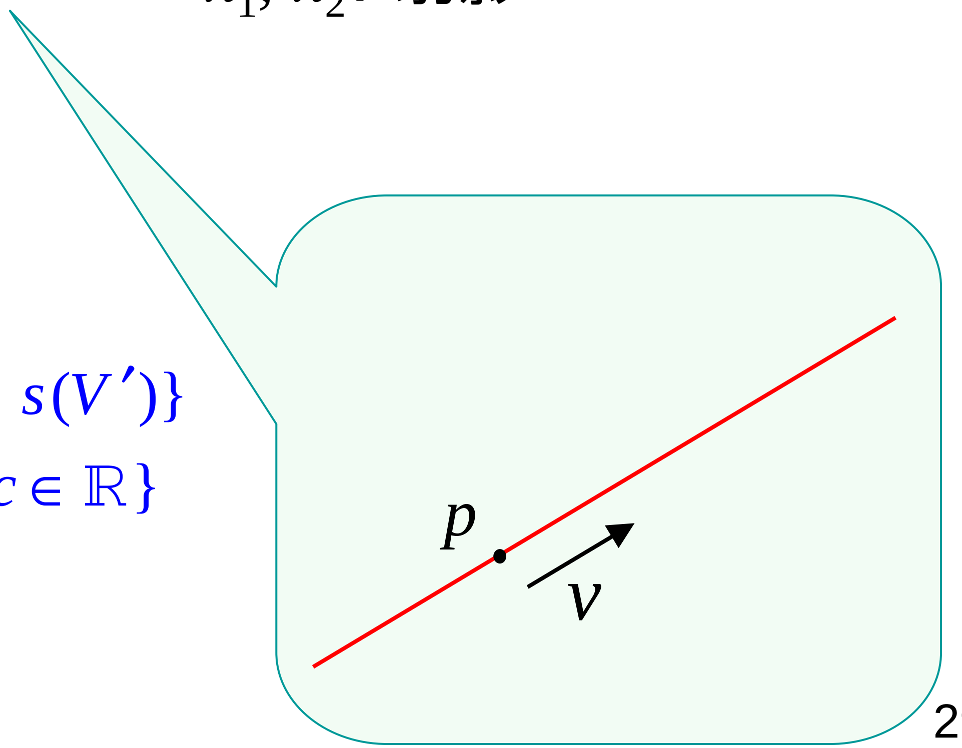
π_1, π_2: 射影

$$s(V\times\mathbb{R}) = \{(v,c) \mid c\in\mathbb{R}\}$$

$$s(V') = \{cv \mid c\in\mathbb{R}\}$$

$$s(X\times X') = \{(x,p) \mid x-p \in s(V')\}$$

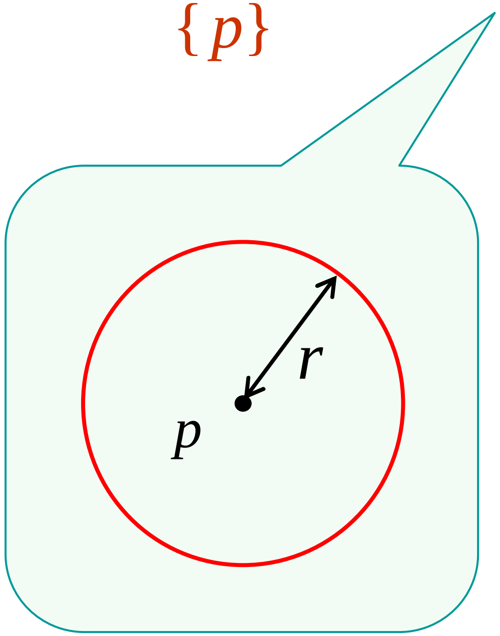
$$s(X) = \{x \mid x = p + cv, c\in\mathbb{R}\}$$

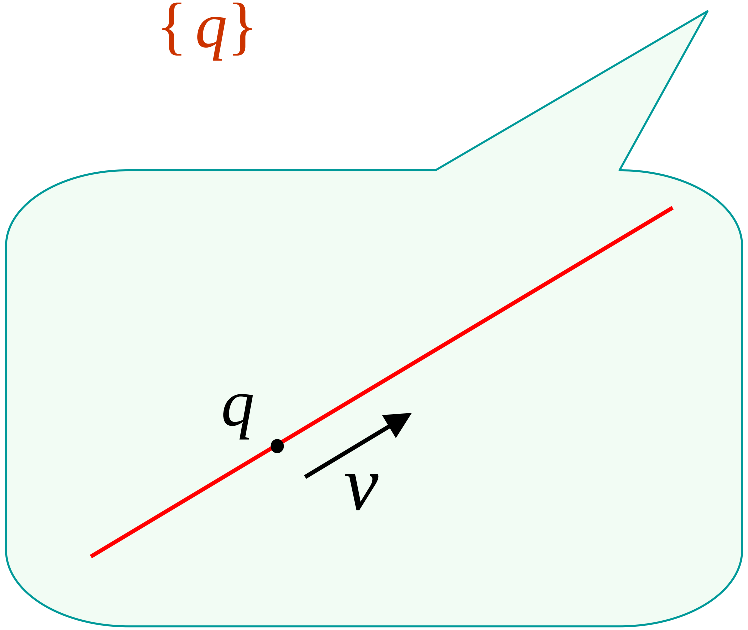


和集合

$$
\begin{tikzcd}
 & \mathbb{R} \arrow[r, dashed, "\mathrm{dist}^{-1}"] & X\times X' \arrow[d, dashed, "\pi_1"] \\
\{r\} \arrow[ur, phantom, "\subset" sloped] & X' \arrow[ur, dashed, "\pi_2^{-1}"] & X \\
\{p\} \arrow[ur, phantom, "\subset" sloped] & &
\end{tikzcd}
$$

$$
\begin{tikzcd}
 & V \arrow[r, dashed, "\pi_1^{-1}"] & V\times\mathbb{R} \arrow[r, dashed, "\mathrm{mult}"] & V' \arrow[dl, dashed, "\mathrm{sub}^{-1}"'] \\
\{v\} \arrow[ur, phantom, "\subset" sloped] & X' \arrow[r, dashed, "\pi_2^{-1}"] & X\times X' \arrow[r, dashed, "\pi_1"] & X \\
\{q\} \arrow[ur, phantom, "\subset" sloped] & & &
\end{tikzcd}
$$

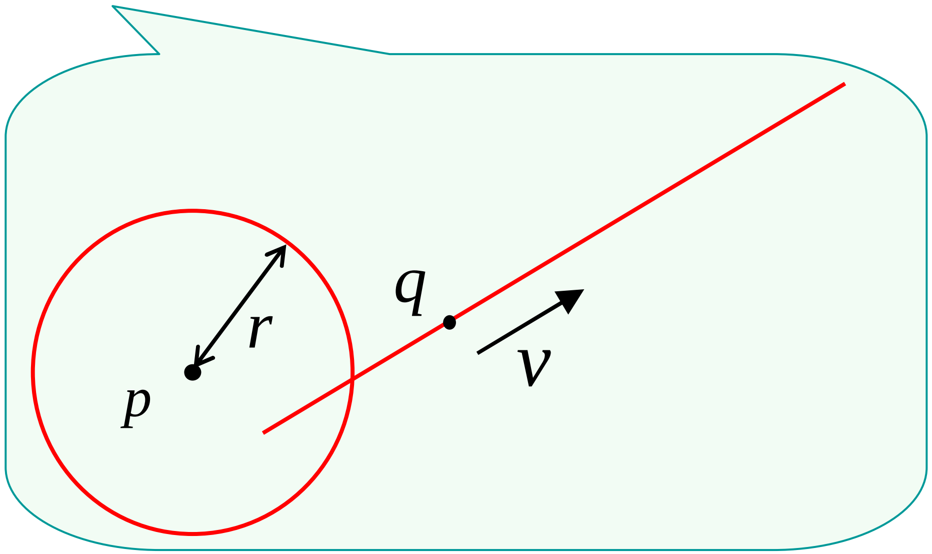




和集合

$$
\begin{tikzcd}
\{r\} \subset \mathbb{R} \arrow[r, dotted, "\mathrm{dist}^{-1}"] & X\times X' \arrow[dd, dotted, "\pi_1"] \\
\{p\} \subset X' \arrow[ur, dotted, "\pi_2^{-1}"] & \\
 & X
\end{tikzcd}
\qquad
\begin{tikzcd}
\{v\} \subset V \arrow[r, dotted, "\pi_1^{-1}"] & V\times\mathbb{R} \arrow[r, dotted, "\mathrm{mult}"] & V' \arrow[dl, dotted, "\mathrm{sub}^{-1}"'] \\
\{q\} \subset X' \arrow[r, dotted, "\pi_2^{-1}"] & X\times X' \arrow[d, dotted, "\pi_1"] & \\
 & X &
\end{tikzcd}
$$

$X \supset$



再帰的定義

$$
\begin{tikzcd}
 & V \arrow[r, dotted, "\pi_2^{-1}"] & X\times V \arrow[d, dotted, shift left, "\mathrm{add}"] & \\
\{v\} \arrow[ur, phantom, "\subset"] & & & \\
 & X' \arrow[r, dotted, "\mathrm{id.}"] & X \arrow[u, dotted, shift left, "\pi_1^{-1}"] \arrow[dr, phantom, "\supset"] & \\
\{p\} \arrow[ur, phantom, "\subset"] & & & ?
\end{tikzcd}
$$

X, X' :ユークリッド平面

V : 2次元ベクトル空間

add: $X\times V \to X \quad (x,v) \mapsto x+v$

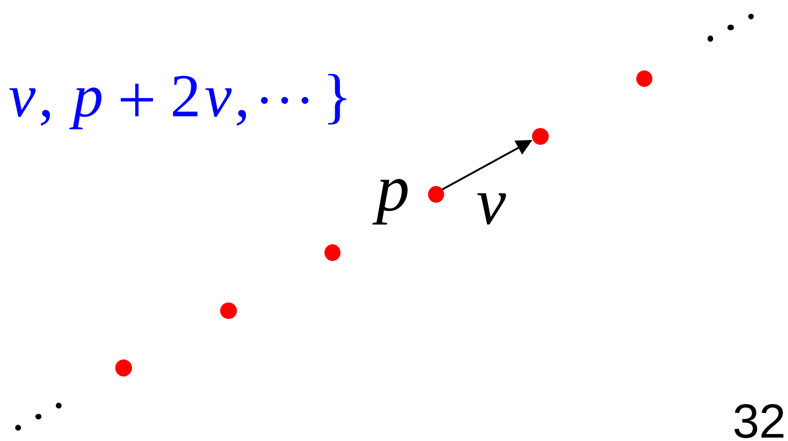
π_1, π_2: 射影

$$s(X\times V) = \{(x,v) \mid x \in s(X)\}$$

$$s(X) = s(X') \cup \mathrm{add}(s(X\times V))$$

$$= \{p\} \cup \{x+v \mid x \in s(X)\}$$

$$\supset \{\cdots, p-2v, p-v, p, p+v, p+2v, \cdots\}$$



再帰的定義

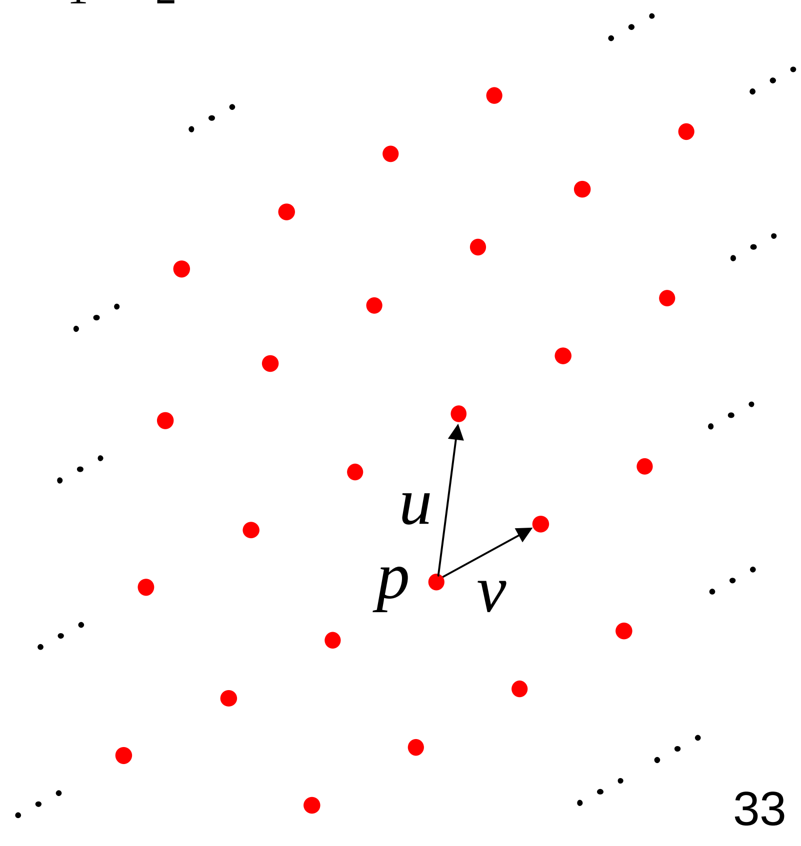
$$
\begin{tikzcd}
V \arrow[r, dotted, "\pi_2^{-1}"] \arrow[d, phantom, "\cup"] & X\times V \arrow[d, dotted, shift left, "\mathrm{add}"] \\
\{u,v\} & X \arrow[u, dotted, shift left, "\pi_1^{-1}"] \arrow[d, phantom, "\cup"] \\
X' \arrow[ur, dotted, hook, "\mathrm{id.}"] \arrow[d, phantom, "\cup"] & ? \\
\{p\} &
\end{tikzcd}
$$

X, X' :ユークリッド平面

V：2次元ベクトル空間

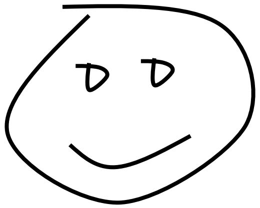
add: $X\times V \rightarrow X \quad (x,v) \mapsto x+v$

π_1, π_2: 射影



階層的定義

$$
\begin{tikzcd}
V \arrow[r, dotted, "\pi_2^{-1}"] \arrow[d, phantom, "\cup"] & X\times V \arrow[d, dotted, shift left, "\mathrm{add}"] \\
\{u,v\} & \\
X' \arrow[r, dotted, hook, "\mathrm{id.}"] \arrow[d, phantom, "\cup"] & X \arrow[uu, dotted, shift left, "\pi_1^{-1}"] \arrow[d, phantom, "\cup"] \\
A & s(X)=?
\end{tikzcd}
$$



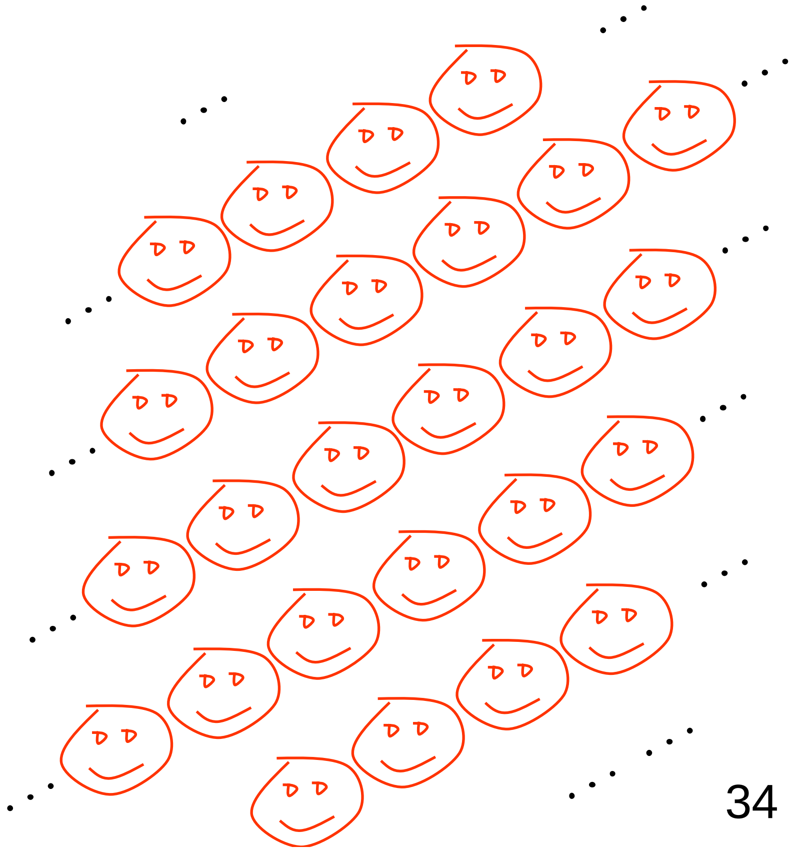
$A \subset X$

X, X' :ユークリッド平面

V：2次元ベクトル空間

add: $X\times V \to X \quad (x,v) \mapsto x+v$

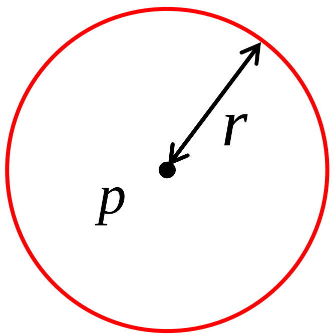
π_1, π_2：射影

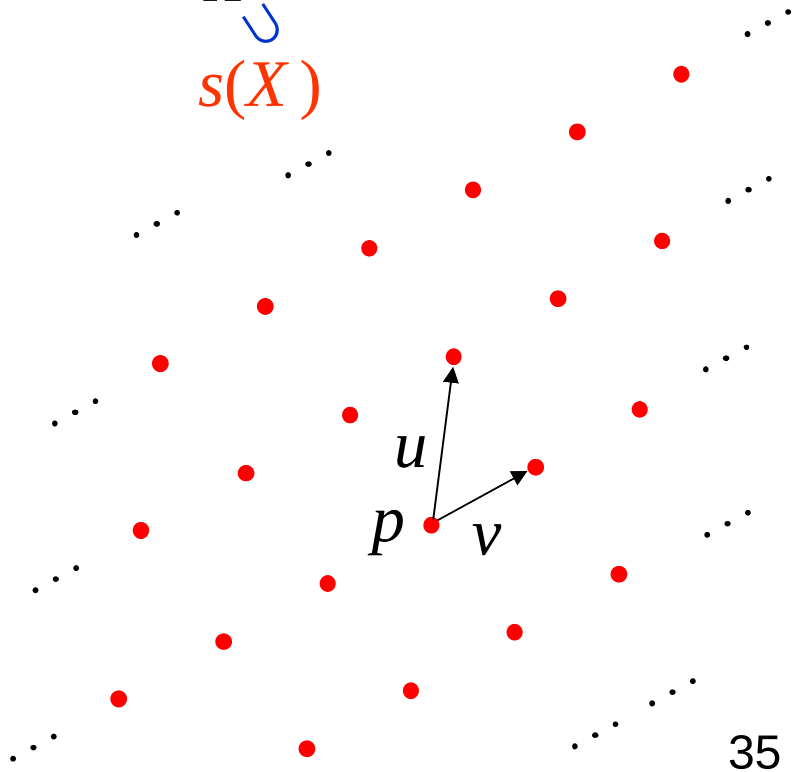


階層的定義

$$
\begin{tikzcd}
\{r\} \arrow[r, phantom, "\subset"] & \mathbb{R} \arrow[r, dotted, "\mathrm{dist}^{-1}"] & X\times X' \arrow[d, dotted, "\pi_1"] \\
\{p\} \arrow[r, phantom, "\subset"] & X' \arrow[ur, dotted, "\pi_2^{-1}"] & X
\end{tikzcd}
$$

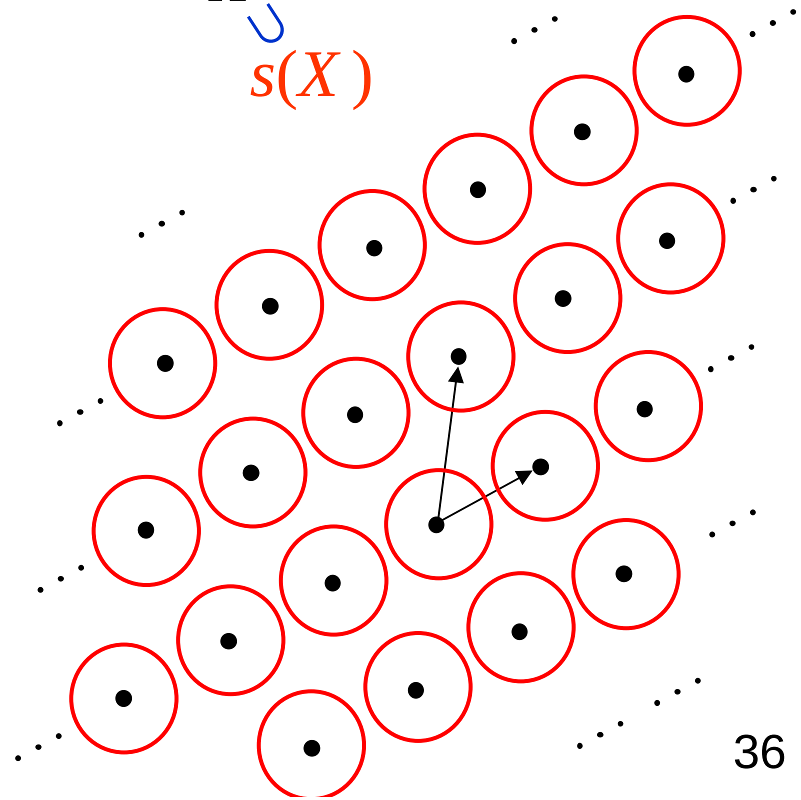
$$
\begin{tikzcd}
\{u,v\} \arrow[r, phantom, "\subset"] & V \arrow[r, dotted, "\pi_2^{-1}"] & X\times V \arrow[d, dotted, hook, bend left, "\mathrm{add}"] \\
\{p\} \arrow[r, phantom, "\subset"] & X' \arrow[r, dotted, hook, "\mathrm{id.}"] & X \arrow[u, dotted, bend left, "\pi_1^{-1}"] \arrow[r, phantom, "\supset"] & s(X)
\end{tikzcd}
$$





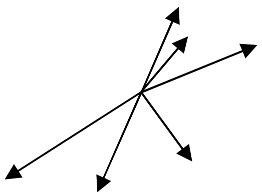
階層的定義

$$
\begin{tikzcd}
 & \mathbb{R} \arrow[r, dotted, "\mathrm{dist}^{-1}"] & X\times X' \arrow[ddrr, dotted, "\pi_1"] & & V \arrow[r, dotted, "\pi_2^{-1}"] & X\times V \arrow[d, dotted, shift left, "\mathrm{add}"] \\
\{r\} \arrow[ur, phantom, "\subset"] & & & \{u,v\} \arrow[ur, phantom, "\subset"] & & \\
 & X' \arrow[uur, dotted, "\pi_2^{-1}"] & & & X' \arrow[r, dotted, hook, "\mathrm{id.}"] & X \arrow[u, dotted, shift left, "\pi_1^{-1}"] \\
\{p\} \arrow[ur, phantom, "\subset"] & & & & & s(X) \arrow[u, phantom, "\subset"]
\end{tikzcd}
$$

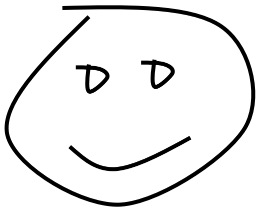


データの規則的部分と不規則部分

- 例: 同じパターンが不規則に繰り返す
  - 規則性: パターンが同じ
  - 不規則性:
    - パターンは不規則でよい
    - パターンがどこに現れるか



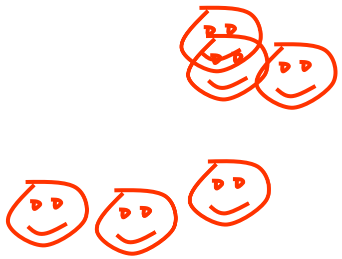
$R\subset V$



$A\subset X$

$$\begin{tikzcd}
R \arrow[r, phantom, "\subset"] & V \arrow[r, dotted, "\pi_2^{-1}"] & X\times V \arrow[d, dotted, "\mathrm{add}"] \\
A \arrow[r, phantom, "\subset"] & X' \arrow[ur, dotted, "\pi_1^{-1}"] & X
\end{tikzcd}$$

$$s(X)=?$$



計算の表現

- 例：階乗の計算

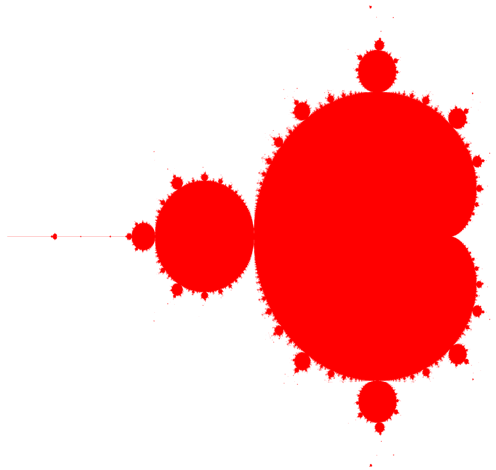
$$\begin{tikzcd}
\mathbb{N}\times\mathbb{N} \arrow[r, dotted, hook] & \mathbb{N}\times\mathbb{N} \arrow[loop right, dotted, "(n,m)\mapsto(n+1,m(n+1))"] \\
\{(0,1)\} \arrow[u, phantom, "\cup" rotate=-45] & \{(0,1),(1,1),(2,2),(3,6),\cdots,(n,n!),\cdots\} \arrow[u, phantom, "\cup" rotate=45]
\end{tikzcd}$$

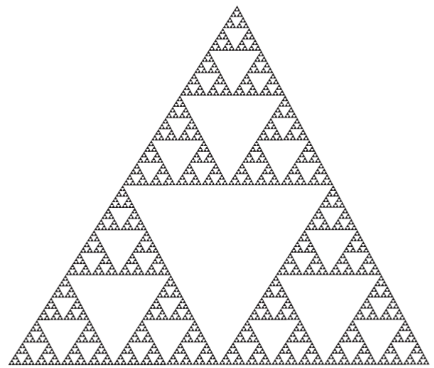
- 例：Fibonacci数の計算

$$\begin{tikzcd}
\mathbb{N}^{+}\times\mathbb{N}^{+} \arrow[r, dotted, hook, "\mathrm{id}"] & \mathbb{N}^{+}\times\mathbb{N}^{+} \arrow[r, dotted, "\pi_1"] \arrow[loop below, dotted, "(n,m)\mapsto(m,n+m)"] & \mathbb{N}^{+} \\
\{(1,1)\} \arrow[u, phantom, "\cup" rotate=-45] & & \{1,1,2,3,5,8,\cdots\} \arrow[u, phantom, "\cup" rotate=45]
\end{tikzcd}$$

- 実は任意のTuring machineをエミュレート可能

計算の表現





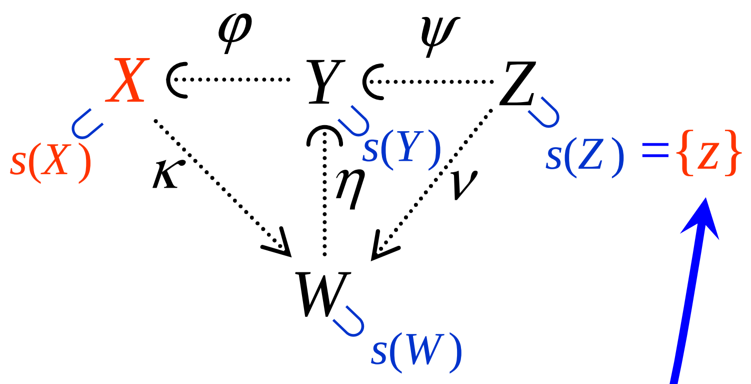
$$
\begin{tikzcd}
\mathbb{C}\times\mathbb{C}\times\mathbb{N} \arrow[loop left, dotted, "{(c,z,k)\mapsto(c,c+z^2,k)}"] \arrow[r, dotted, "\pi_1\times\pi_2"] & \mathbb{C}\times\mathbb{C} \arrow[r, dotted, "\mathrm{cmpl}\circ\pi_1"] & \color{red}{\mathbb{C}} \\
\mathbb{C}\times\mathbb{C}\times\mathbb{N} \arrow[u, hook, dotted, "\mathrm{id}"] & \mathbb{C} \arrow[u, hook, dotted, "\pi_2^{-1}"] & \\
\mathbb{C}\times\{0\}\times\{0\} \arrow[u, phantom, sloped, "\subset"] & \{z\in\mathbb{C} \mid |z|>2\} \arrow[u, phantom, sloped, "\subset"] &
\end{tikzcd}
$$

情報計量

- 任意の部分集合$A \subset X$の情報計量を定義したい
- 写像の集合$\mathcal{M}$を固定し、その元の組み合わせにより作ることが可能な写像
  - id: $X \to X$, π_i: $X_1 \times \ldots \times X_n \to X_i$, $\omega : X \to \{0\}$
  - $f : X \to Y, g : Y \to Z \Rightarrow g \circ f : X \to Z$
  - $f_i : X \to Y_i \Rightarrow f_1 \times \ldots \times f_n : X \to Y_1 \times \ldots \times Y_n$
- それらだけを使ってできる図式($\mathcal{M}$で生成される図式)のうち対象を表現可能なものの最小サイズで$I(A|\mathcal{M})$を定義

図式と断面による表現

- 集合とそれらの間の冪写像を指定(図式)



- いくつかの $s(\,)$ を指定する (部分断面)
- 図式内の集合を一つ指定する

情報計量

問題点

何の制限もしないと、部分断面としてAを指定することで図式のサイズは 0 にできてしまう

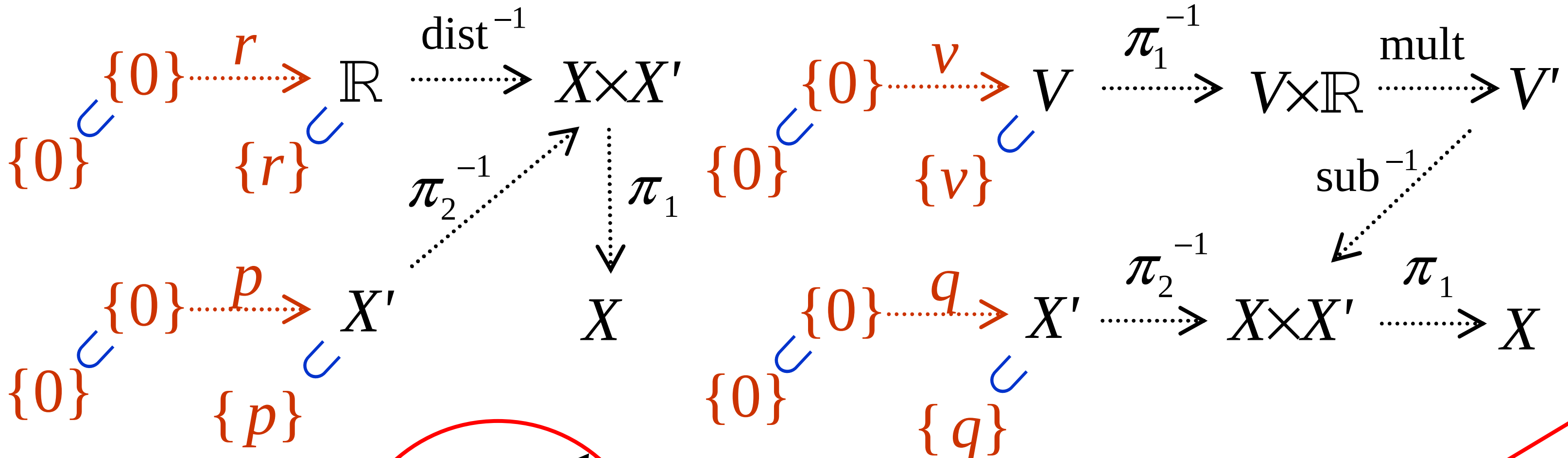
$$s(X)=A \subset X$$

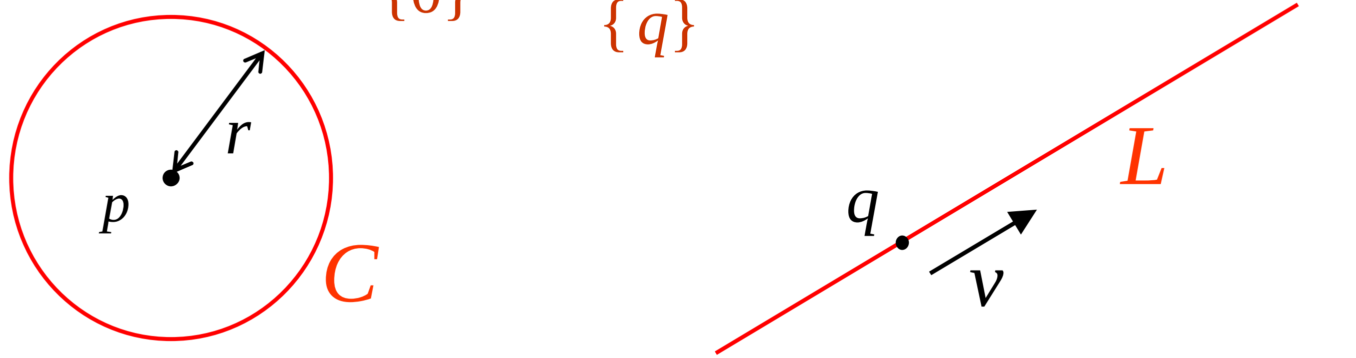
自明な表現

情報計量

定数 $x \in X$ を写像 $x:\{0\}\to X$ $\;(x(0)=x)$とみなし、

部分断面は $s(\{0\})=\{0\}$ のみを許す





$$I(C\mid\{r,p,\mathrm{dist}\})\le 5 \qquad I(L\mid\{v,q,\mathrm{mult},\ \mathrm{sub}\})\le 7$$

情報計量

対象の属する空間を特徴づける写像の組を選び
それに相対的に情報計量を定義する

たとえばユークリッド空間 X の場合、

$$\mathcal{M}_{\mathrm{E}} = \{\mathrm{dist, add, sub, mult}\} \cup X \cup V \cup \mathbb{R}$$

$$\{x{:}\{0\} \to X \mid x \in X\}$$

とすると、

- $I(C \mid \mathcal{M}_{\mathrm{E}}) \leq 5,\ I(L \mid \mathcal{M}_{\mathrm{E}}) \leq 7$
- 任意の有限集合 $A \subset X$ について $I(A|\mathcal{M}_{\mathrm{E}}) \leq |A|$

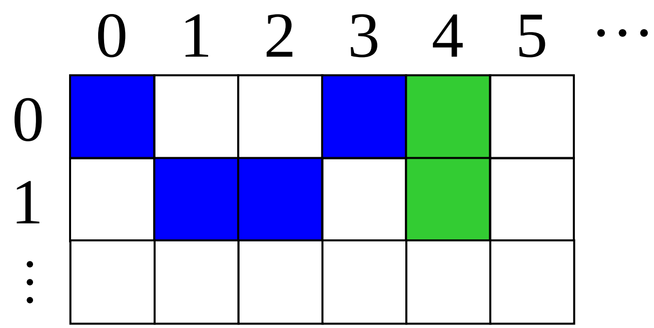
（有限個の点の集合の情報量は点の数で押さえられる）

Kolmogorov複雑性との関係

$\sigma \in b^*$ に対して$\bar{\sigma} \subset \mathbb{N}\times\mathbb{N}$を次のように定義する

$$\bar{\sigma} = \{(i,\sigma[i]) \mid i = 0, 1,\cdots,|\sigma|-1\} \cup \{(|\sigma|,0), (|\sigma|,1)\}$$

例:$\sigma=0110$なら$\bar{\sigma}=\{(0,0),(1,1),(2,1),(3,0),(4,0),(4,1)\}$

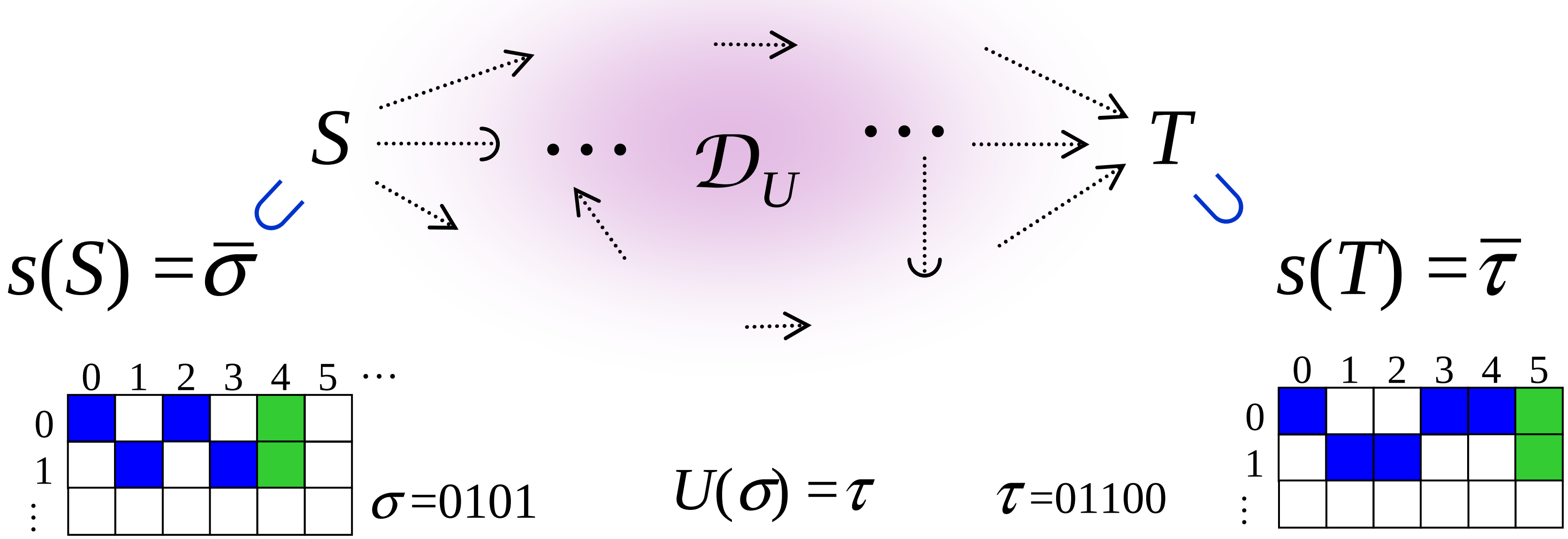


自然数を特徴づける写像：$\mathcal{M}_{\mathbb{N}}=\{0, \mathrm{succ}\}$

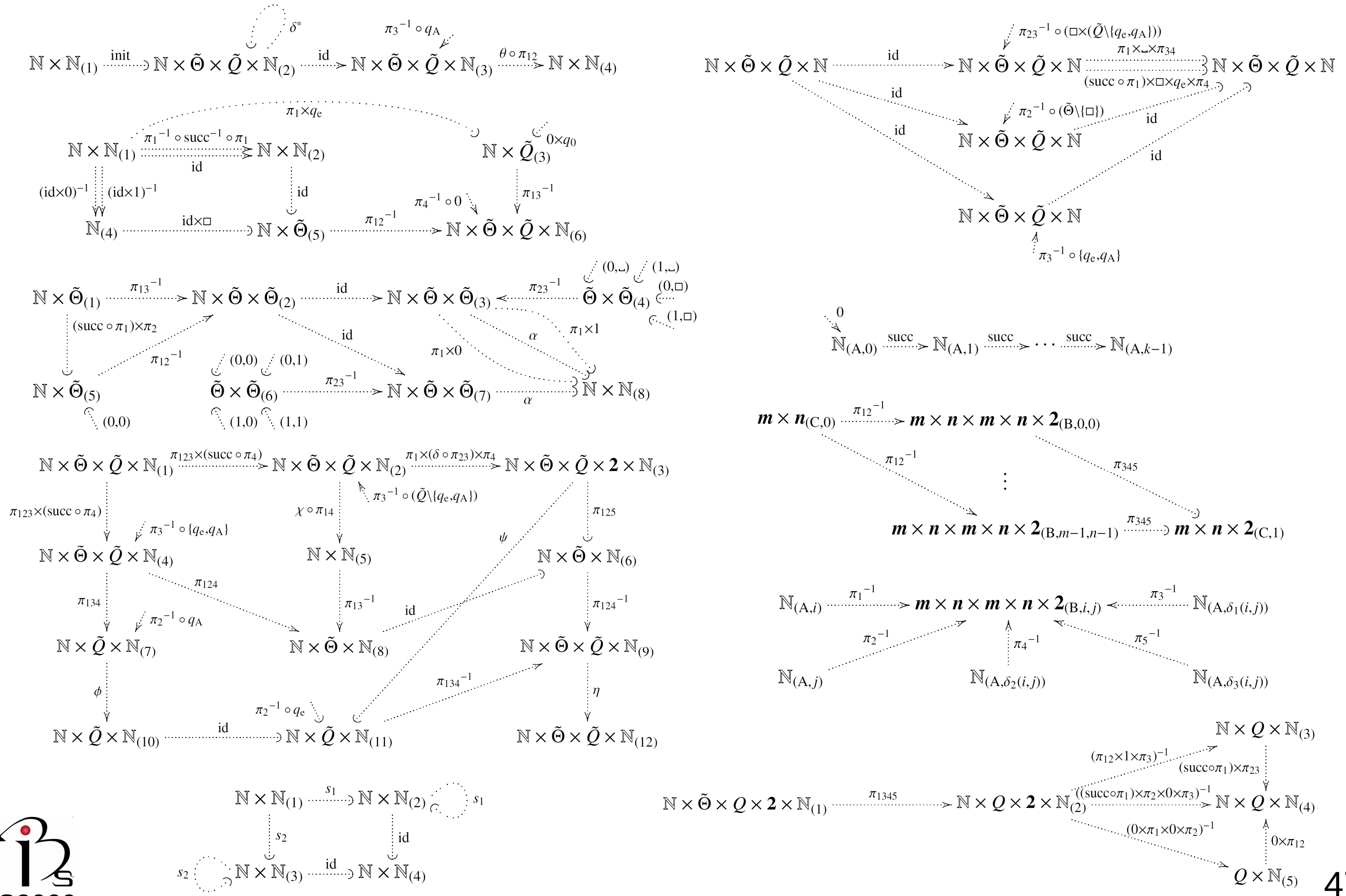
- $0:\{0\} \to \mathbb{N},\ 0(0) = 0$
- $\mathrm{succ}: \mathbb{N} \to \mathbb{N}\quad [\mathrm{succ}(n) = n+1]$

Kolmogorov複雑性との関係

定理 任意の万能 Turing Machine Uについて、$\mathcal{M}_{\mathbb{N}}$ で生成される図式$\mathcal{D}_U$で、$S = \mathbb{N}\times\mathbb{N}$, $T = \mathbb{N}\times\mathbb{N}$ を含み、任意の$\sigma \in b^*$と $s(S) = \overline{\sigma}$なる任意の断面 s について、$U(\sigma) = \uparrow$ なら $s(T) = \varnothing$, $U(\sigma) = \tau \in b^*$ なら $s(T) = \overline{\tau}$ を満たすものが存在する。



Kolmogorov複雑性との関係



IBIS2009

Kolmogorov複雑性との関係

系 任意の万能 Turing Machine Uについて、
定数c_U が存在して、任意の$\sigma \in b^*$について

$$I(\overline{\sigma} | \mathcal{M}_{\mathbb{N}}) \leq 6K_U(\sigma) + c_U$$

を満たす。

逆方向も証明できる：

定理 任意の万能 Turing Machine Uについて

$$K_U(\sigma) \leq d_U I(\sigma | \mathcal{M}_{\mathbb{N}}) + e_U$$

を任意の$\sigma \in b^*$ について満たす定数d_U, e_U が
存在する。

プロセス

一般の対象

近似
符号化

ボトルネック 記号列

アルゴリズム
による圧縮

符号

抽象

符号

一般の対象

表現
圧縮

非記号計算による表現

近似
符号化

符号

IBIS2009

結論

- 非記号データを含む一般の対象の表現
- 非記号の対象について直接計算を定義

  → 非記号計算　（接地計算）
- 非記号の対象の圧縮を直接定義することによる一般の対象についての情報計量の定義
- パターン発見の定式化