Complexity Monotone in Conditions and Future Prediction Errors

Alexey Chernov Marcus Hutter

Istituto Dalle Molle di Studio sull'Intelligenza Artificiale
Lugano, Switzerland

Supported by SNF grants 200020-107590/1 (to Jürgen Schmidhuber),
2100-67712 and 200020-107616

Question

$$\log_2 \frac{\mu(y)}{M(y)} \stackrel{+}{\leq} KP(\mu) \qquad \log_2 \frac{\mu(y|x)}{M(y|x)} \stackrel{+}{\leq} KP(\mu|x) \ ?$$

μ computable measure
M a priori (universal) semicomputable semimeasure
KP prefix complexity

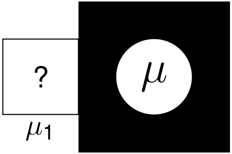
Question

$$\log_2 \frac{\mu(y)}{M(y)} \overset{+}{\leq} KP(\mu)$$

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \overset{+}{\leq} KP(\mu|x) \ ?$$

μ computable measure
M a priori (universal) semicomputable semimeasure
KP prefix complexity

Sequence

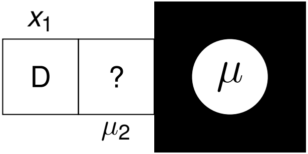


finite alphabet $\mathcal{X} \ni x_1, x_2, \ldots$

μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot | x_1 \ldots x_i) = \frac{\mu(x_1 \ldots x_i \cdot)}{\mu(x_1 \ldots x_i)}$$

Sequence

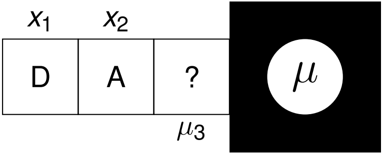


finite alphabet $\mathcal{X} \ni x_1, x_2, \ldots$

μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot|x_1 \ldots x_i) = \frac{\mu(x_1 \ldots x_i \cdot)}{\mu(x_1 \ldots x_i)}$$

Sequence

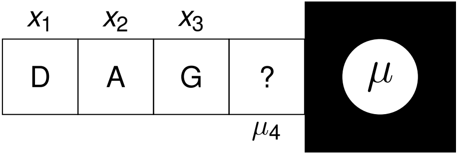


finite alphabet $\mathcal{X} \ni x_1, x_2, \dots$

μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot | x_1 \dots x_i) = \frac{\mu(x_1 \dots x_i \cdot)}{\mu(x_1 \dots x_i)}$$

Sequence

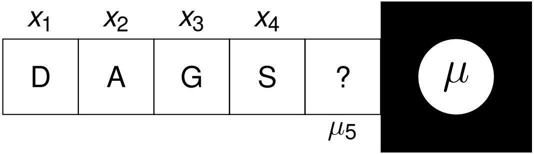


finite alphabet $\mathcal{X} \ni x_1, x_2, \ldots$

μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot|x_1 \ldots x_i) = \frac{\mu(x_1 \ldots x_i \cdot)}{\mu(x_1 \ldots x_i)}$$

Sequence

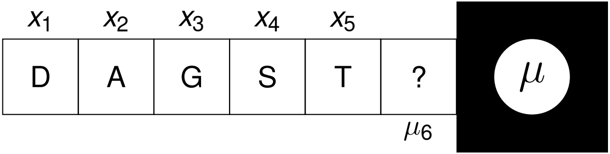


finite alphabet $\mathcal{X} \ni x_1, x_2, \ldots$

μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot|x_1 \ldots x_i) = \frac{\mu(x_1 \ldots x_i \cdot)}{\mu(x_1 \ldots x_i)}$$

Sequence

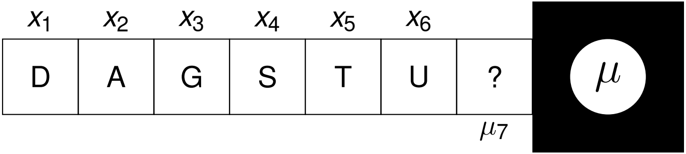


finite alphabet $\mathcal{X} \ni x_1, x_2, \ldots$

μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot|x_1 \ldots x_i) = \frac{\mu(x_1 \ldots x_i \cdot)}{\mu(x_1 \ldots x_i)}$$

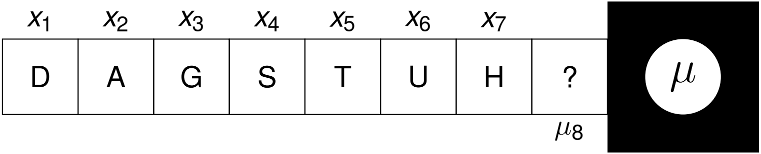
Sequence



finite alphabet $\mathcal{X} \ni x_1, x_2, \ldots$ μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot|x_1 \ldots x_i) = \frac{\mu(x_1 \ldots x_i \cdot)}{\mu(x_1 \ldots x_i)}$$

Sequence



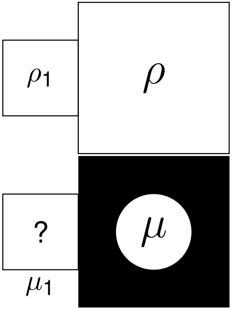
finite alphabet $\mathcal{X} \ni x_1, x_2, \ldots$

μ is a measure

$$\mu_{i+1}(\cdot) = \mu(\cdot | x_1 \ldots x_i) = \frac{\mu(x_1 \ldots x_i \cdot)}{\mu(x_1 \ldots x_i)}$$

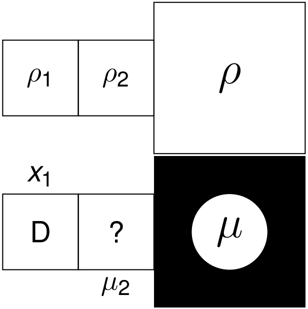
Predictor

Predictor ρ: $x_1, \ldots, x_i \quad \mapsto \quad \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$



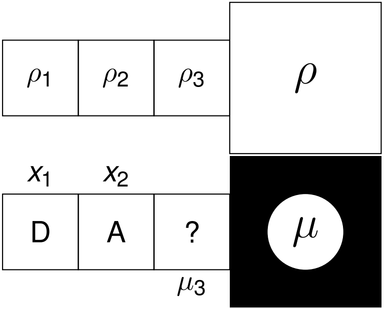
Predictor

Predictor ρ: $x_1, \ldots, x_i \quad \mapsto \quad \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$



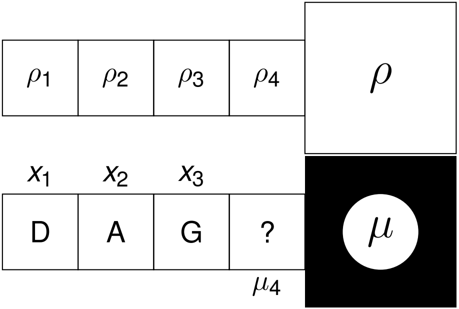
Predictor

Predictor ρ: $x_1, \ldots, x_i \mapsto \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$



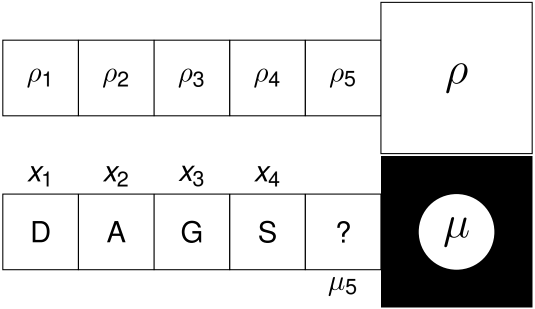
Predictor

Predictor ρ: $x_1, \ldots, x_i \quad \mapsto \quad \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$



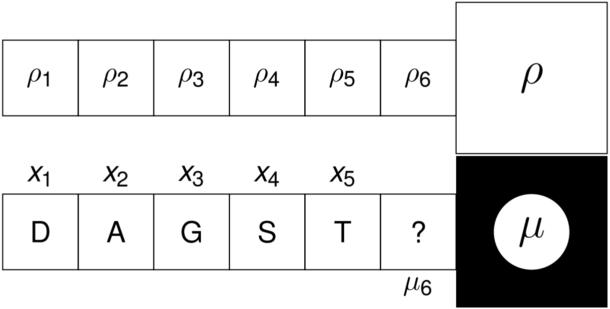
Predictor

Predictor ρ: $x_1, \ldots, x_i \quad \mapsto \quad \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$



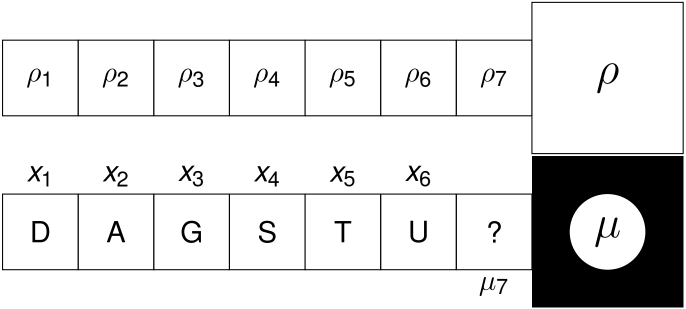
Predictor

Predictor ρ: $$x_1, \ldots, x_i \quad \mapsto \quad \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$$



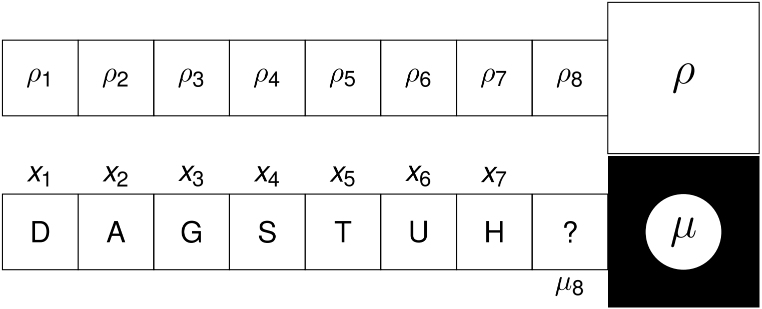
Predictor

Predictor ρ: $x_1, \ldots, x_i \quad \mapsto \quad \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$



Predictor

Predictor ρ: $$x_1, \ldots, x_i \quad \mapsto \quad \rho_{i+1}(\cdot) \approx \mu_{i+1}(\cdot)$$



Quality of prediction

$\rho_i(\cdot) \approx \mu_i(\cdot)$ with high μ-probability:

$$\mathrm{Dist}(\rho,\mu) = \mathbf{E}\sum_{i=1}^{\infty} \mathrm{dist}_{x_1\ldots x_i}(\rho_i,\mu_i) \quad = \sum_{x_1 x_2\ldots} \mu(x_1 x_2\ldots) \sum_{i=1}^{\infty} \mathrm{dist}_{x_1\ldots x_i}(\rho_i,\mu_i)$$

$$\mathrm{dist}_{x_1\ldots x_i}(\rho_i,\mu_i) = \frac{1}{\ln 2} \times$$

$\sum_{a\in\mathcal{X}} (\rho_i(a) - \mu_i(a))^2$ or $\frac{1}{2}\left(\sum_{a\in\mathcal{X}} |\rho_i(a) - \mu_i(a)|\right)^2$ or

$\sum_{a\in\mathcal{X}} \left(\sqrt{\rho_i(a)} - \sqrt{\mu_i(a)}\right)^2$ or $\sum_{a\in\mathcal{X}} \mu_i(a) \ln \frac{\mu_i(a)}{\rho_i(a)}$

$$0 \le \mathrm{Dist}(\rho,\mu) \le D_\rho := \mathbf{E}\log_2 \frac{\mu(x_1 x_2 \ldots)}{\rho(x_1 x_2 \ldots)}$$

Intuitively: (for a deterministic μ)

$D_\rho \sim$ number of prediction errors

Solomonoff prior

$$\rho_i(\cdot) = \frac{M(x_1 \dots x_i \cdot)}{M(x_1 \dots x_i)} \qquad M(x) = \sum_\mu w_\mu \mu(x)$$

M is a Bayes mixture of all semi-computable semi-measures.

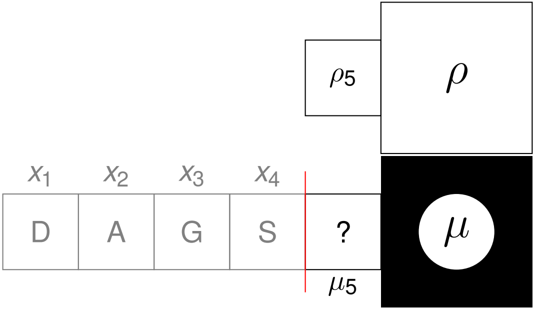
Theorem (Solomonoff 1964, 1978)

For any computable measure μ

$$\mathrm{Dist}(M, \mu) \leq D_M \stackrel{+}{\leq} KP(\mu)$$

$KP(\mu)$ is Kolmogorov complexity of μ
$\sim$ quantity of information in μ
$\sim$ the size of the shortest description of μ

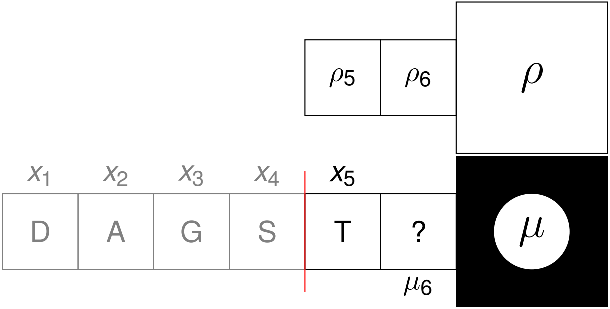
Prediction with a priori information



$x_1, \ldots, x_n$ are fixed

$$\mathrm{Dist}(\rho, \mu | x_1 \ldots x_n) = \mathop{\mathbf{E}}_{x_{n+1} x_{n+2} \ldots} \sum_{i=n+1}^{\infty} \mathrm{dist}_{x_1 \ldots x_i}(\rho_i, \mu_i)$$

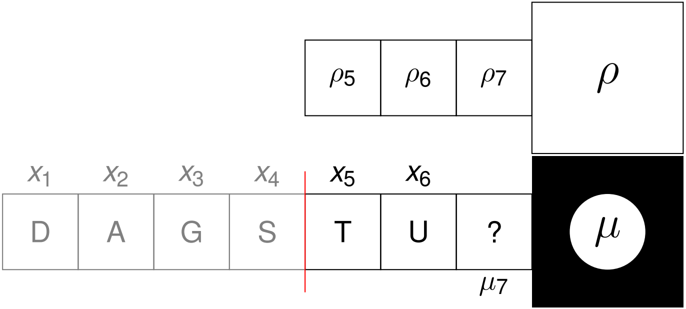
Prediction with a priori information



$x_1, \ldots, x_n$ are fixed

$$\mathrm{Dist}(\rho, \mu | x_1 \ldots x_n) = \mathop{\mathbf{E}}_{x_{n+1} x_{n+2} \ldots} \sum_{i=n+1}^{\infty} \mathrm{dist}_{x_1 \ldots x_i}(\rho_i, \mu_i)$$

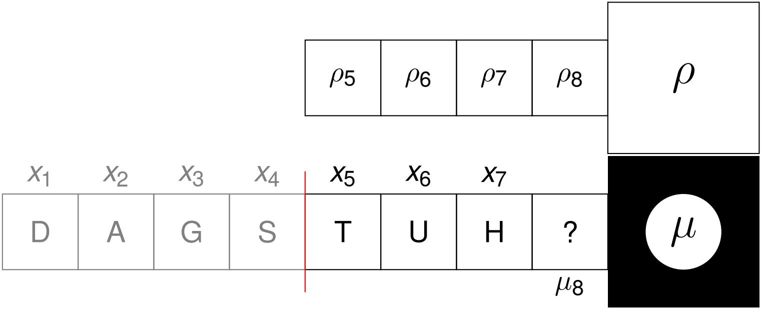
Prediction with a priori information



$x_1, \ldots, x_n$ are fixed

$$\mathrm{Dist}(\rho, \mu | x_1 \ldots x_n) = \mathop{\mathbf{E}}_{x_{n+1} x_{n+2} \ldots} \sum_{i=n+1}^{\infty} \mathrm{dist}_{x_1 \ldots x_i}(\rho_i, \mu_i)$$

Prediction with a priori information



$x_1, \ldots, x_n$ are fixed

$$\mathrm{Dist}(\rho, \mu | x_1 \ldots x_n) = \mathop{\mathbf{E}}_{x_{n+1} x_{n+2} \ldots} \sum_{i=n+1}^{\infty} \mathrm{dist}_{x_1 \ldots x_i}(\rho_i, \mu_i)$$

The problem

$x = x_1 \dots x_n \qquad \mathrm{Dist}(M, \mu | x) \leq D_M(x) := \mathbf{E}_y \log_2 \frac{\mu(y_1 y_2 \dots | x)}{M(y_1 y_2 \dots | x)}$

For any computable measure μ, for any word x

$$\frac{\mu(y|x)}{M(y|x)} \leq \quad ?$$

We know

We want

$$\log_2 \frac{\mu(y)}{M(y)} \overset{+}{\leq} KP(\mu)$$

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \overset{+}{\leq} KP(\mu|x)$$

If x contains a lot of information about μ ($KP(\mu|x)$ is small), prediction is easy.

The problem

$$x = x_1 \dots x_n \qquad \mathrm{Dist}(M, \mu|x) \leq D_M(x) := \mathbf{E}_y \log_2 \frac{\mu(y_1 y_2 \dots |x)}{M(y_1 y_2 \dots |x)}$$

For any computable measure μ, for any word x

$$\frac{\mu(y|x)}{M(y|x)} \leq \quad ?$$

We know

$$\log_2 \frac{\mu(y)}{M(y)} \stackrel{+}{\leq} KP(\mu)$$

We want

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \stackrel{+}{\leq} KP(\mu|x)$$

If x contains a lot of information about μ ($KP(\mu|x)$ is small), prediction is easy.

$KP(\mu|x)$ bound

Theorem

For any computable measure μ and any $x, y \in \mathcal{X}^*$

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \overset{+}{\leq} KP(\mu|x) + KP(\ell(x))$$

Corollary

1. $\mathrm{Dist}(M, \mu|x_1 \dots x_n) \overset{+}{\leq} KP(\mu|x_1 \dots x_n) + KP(n)$
2. $\mathrm{Dist}(M, \mu) \overset{+}{\leq} \min_n \{ \mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \frac{2}{\ln 2} n \}$

$KP(\mu|x)$ bound

Theorem

For any computable measure μ and any $x, y \in \mathcal{X}^*$

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \stackrel{+}{\leq} KP(\mu|x) + KP(\ell(x))$$

Corollary

1. $\mathrm{Dist}(M, \mu|x_1 \ldots x_n) \stackrel{+}{\leq} KP(\mu|x_1 \ldots x_n) + KP(n)$
2. $\mathrm{Dist}(M, \mu) \stackrel{+}{\leq} \min_n \{\mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \frac{2}{\ln 2} n\}$

Example

$$\mathrm{Dist}(M,\mu) \overset{+}{\leq} \min_n\{\mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \tfrac{2}{\ln 2} n\}$$

x_1

?

"number of errors" $\sim \mathrm{Dist}(M,\mu)$

Solomonoff bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu) \sim \text{"size of the image"} \approx 10^5$$

New bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu|x_1) + KP(1) + \tfrac{2}{\ln 2} \sim \text{"small constant"}$$

Example

$$\mathrm{Dist}(M,\mu) \overset{+}{\leq} \min_n \{\mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \frac{2}{\ln 2} n\}$$

x_1 x_2

"number of errors" $\sim \mathrm{Dist}(M,\mu)$

Solomonoff bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu) \sim \text{“size of the image”} \approx 10^5$$

New bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu|x_1) + KP(1) + \frac{2}{\ln 2} \sim \text{“small constant”}$$

Example

$$\mathrm{Dist}(M,\mu) \overset{+}{\leq} \min_n \{ \mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \tfrac{2}{\ln 2} n \}$$

x_1

x_2

x_3

?

"number of errors" $\sim \mathrm{Dist}(M,\mu)$

Solomonoff bound:

$\mathrm{Dist}(M,\mu) \lesssim KP(\mu) \sim$ "size of the image" $\approx 10^5$

New bound:

$\mathrm{Dist}(M,\mu) \lesssim KP(\mu|x_1) + KP(1) + \frac{2}{\ln 2} \sim$ "small constant"

Example

$$\mathrm{Dist}(M, \mu) \overset{+}{\leq} \min_n \{ \mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \frac{2}{\ln 2} n \}$$

x_1

x_2

x_3

x_4

?

"number of errors" $\sim \mathrm{Dist}(M, \mu)$

Solomonoff bound:

$$\mathrm{Dist}(M, \mu) \lesssim KP(\mu) \sim \text{"size of the image"} \approx 10^5$$

New bound:

$$\mathrm{Dist}(M, \mu) \lesssim KP(\mu|x_1) + KP(1) + \frac{2}{\ln 2} \sim \text{"small constant"}$$

Example

$$\mathrm{Dist}(M,\mu) \overset{+}{\leq} \min_n \{ \mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \tfrac{2}{\ln 2} n \}$$

x_1	x_2	x_3	x_4	x_5	
					?

"number of errors" $\sim \mathrm{Dist}(M,\mu)$

Solomonoff bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu) \sim \text{“size of the image”} \approx 10^5$$

New bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu|x_1) + KP(1) + \tfrac{2}{\ln 2} \sim \text{“small constant”}$$

Example

$$\mathrm{Dist}(M,\mu) \overset{+}{\leq} \min_n\{\mathbf{E}_{\ell(x)=n} KP(\mu|x) + KP(n) + \frac{2}{\ln 2} n\}$$

x_1 x_2 x_3 x_4 x_5





?

"number of errors" $\sim \mathrm{Dist}(M,\mu)$

Solomonoff bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu) \sim \text{“size of the image”} \approx 10^5$$

New bound:

$$\mathrm{Dist}(M,\mu) \lesssim KP(\mu|x_1) + KP(1) + \frac{2}{\ln 2} \sim \text{“small constant”}$$

Analysis of the bound: disinformation

$$\mathrm{Dist}(M,\mu|x) \overset{+}{\leq} KP(\mu|x) + KP(\ell(x)) \xrightarrow{\ell(x)\to\infty} \infty$$

Is it worse than $\mathrm{Dist}(M,\mu) \overset{+}{\leq} KP(\mu)$? No!

$$\mathrm{Dist}(M,\mu) = \ldots + \sum_{\ell(x)=n} \mu(x)\mathrm{Dist}(M,\mu|x)$$

$\mu(x) \to 0 \quad \Rightarrow \quad \mathrm{Dist}(M,\mu|x) \to \infty$ is possible

If x is not typical for μ ($\mu(x) \approx 0$), then we get disinformation

Analysis of the bound: disinformation

$$\mathrm{Dist}(M,\mu|x) \overset{+}{\leq} KP(\mu|x) + KP(\ell(x)) \xrightarrow{\ell(x)\to\infty} \infty$$

Is it worse than $\mathrm{Dist}(M,\mu) \overset{+}{\leq} KP(\mu)$? No!

$$\mathrm{Dist}(M,\mu) = \ldots + \sum_{\ell(x)=n} \mu(x)\mathrm{Dist}(M,\mu|x)$$

$\mu(x) \to 0 \quad \Rightarrow \quad \mathrm{Dist}(M,\mu|x) \to \infty$ is possible

If x is not typical for μ ($\mu(x) \approx 0$), then we get disinformation

Analysis of the bound: disinformation

$$\mathrm{Dist}(M,\mu|x) \overset{+}{\leq} KP(\mu|x) + KP(\ell(x)) \xrightarrow{\ell(x)\to\infty} \infty$$

Is it worse than $\mathrm{Dist}(M,\mu) \overset{+}{\leq} KP(\mu)$? No!

$$\mathrm{Dist}(M,\mu) = \ldots + \sum_{\ell(x)=n} \mu(x)\mathrm{Dist}(M,\mu|x)$$

$\mu(x) \to 0 \quad \Rightarrow \quad \mathrm{Dist}(M,\mu|x) \to \infty$ is possible

If x is not typical for μ ($\mu(x) \approx 0$), then we get disinformation

How to deal with disinformation

$$\log_2 \frac{\mu(y|x)}{M(y|x)} = \log_2 \frac{\mu(y)}{M(y)} + \log_2 \frac{M(x)}{\mu(x)} \stackrel{+}{\leq} KP(\mu) + \log_2 \frac{M(x)}{\mu(x)}$$

Randomness deficiency: $d_\mu(x) = \log_2 \frac{M(x)}{\mu(x)}$

d_μ is a measure of non-typicalness, $d_\mu(x)$ is small for most x

$d_\mu(x) = \ell(x) - KP(x)$ for uniform μ

How to deal with disinformation

$$\log_2 \frac{\mu(y|x)}{M(y|x)} = \log_2 \frac{\mu(y)}{M(y)} + \log_2 \frac{M(x)}{\mu(x)} \overset{+}{\leq} KP(\mu) + d_\mu(x)$$

Randomness deficiency: $d_\mu(x) = \log_2 \frac{M(x)}{\mu(x)}$

d_μ is a measure of non-typicalness, $d_\mu(x)$ is small for most x

$d_\mu(x) = \ell(x) - KP(x)$ for uniform μ

How to deal with disinformation

Theorem (An. Muchnik, A. Shen)

For any computable measure μ *and any* $x, y \in \mathcal{X}^*$

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \equiv d_\mu(x) - d_\mu(xy) \overset{+}{\leq} KP(\mu) + KP(\lceil d_\mu(x) \rceil)$$

Randomness deficiency: $d_\mu(x) = \log_2 \frac{M(x)}{\mu(x)}$

d_μ is a measure of non-typicalness, $d_\mu(x)$ is small for most x

$d_\mu(x) = \ell(x) - KP(x)$ for uniform μ

Structure of the bounds

prediction errors	$\Leftarrow$	information	$+$	quantity of disinformation
$\mathrm{Dist}(M,\mu)$	$\stackrel{+}{\leq}$	$KP(\mu)$		
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP(\mu\|x)$	$+$	$KP(\ell(x))$
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP(\mu)$	$+$	$KP(d_\mu(x))$
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP_*(\mu\|x*)$	$+$	$KP(d_\mu(x))$

Structure of the bounds

prediction errors	$\Leftarrow$	information	$+$	quantity of disinformation
$\mathrm{Dist}(M,\mu)$	$\stackrel{+}{\leq}$	$KP(\mu)$		
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP(\mu\|x)$	$+$	$KP(\ell(x))$
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP(\mu)$	$+$	$KP(d_\mu(x))$
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP_*(\mu\|x*)$	$+$	$KP(d_\mu(x))$

Structure of the bounds

prediction errors	$\Leftarrow$	information	$+$	quantity of disinformation
$\mathrm{Dist}(M, \mu)$	$\stackrel{+}{\leq}$	$KP(\mu)$		
$\mathrm{Dist}(M, \mu \vert x)$	$\stackrel{+}{\leq}$	$KP(\mu \vert x)$	$+$	$KP(\ell(x))$
$\mathrm{Dist}(M, \mu \vert x)$	$\stackrel{+}{\leq}$	$KP(\mu)$	$+$	$KP(d_\mu(x))$
$\mathrm{Dist}(M, \mu \vert x)$	$\stackrel{+}{\leq}$	$KP_*(\mu \vert x*)$	$+$	$KP(d_\mu(x))$

Structure of the bounds

prediction errors	$\Leftarrow$	information	$+$	quantity of disinformation
$\mathrm{Dist}(M,\mu)$	$\stackrel{+}{\leq}$	$KP(\mu)$		
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP(\mu\|x)$	$+$	$KP(\ell(x))$
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP(\mu)$	$+$	$KP(d_\mu(x))$
$\mathrm{Dist}(M,\mu\|x)$	$\stackrel{+}{\leq}$	$KP_*(\mu\|x*)$	$+$	$KP(d_\mu(x))$

Structure of the bounds

prediction errors	$\Leftarrow$	information	$+$	quantity of disinformation
$\mathrm{Dist}(M,\mu)$	$\overset{+}{\leq}$	$KP(\mu)$		
$\mathrm{Dist}(M,\mu\vert x)$	$\overset{+}{\leq}$	$KP(\mu\vert x)$	$+$	$KP(\ell(x))$
$\mathrm{Dist}(M,\mu\vert x)$	$\overset{+}{\leq}$	$KP(\mu)$	$+$	$KP(d_\mu(x))$
$\mathrm{Dist}(M,\mu\vert x)$	$\overset{+}{\leq}$	$KP_*(\mu\vert x*)$	$+$	$KP(d_\mu(x))$

Prefix complexity monotone in conditions

Definition (informal)

$x, y \in \mathcal{X}^*$, U is a universal twice-prefix machine, $p \in \{0,1\}^*$

$$KP_*(y|x*) = \min\{\ell(p) | U(p*, x*) = y\}$$

Recall: conditional prefix complexity

$x, y \in \mathcal{X}^*$, U is a universal prefix machine, $p \in \{0,1\}^*$

$$KP(y|x) = \min\{\ell(p) | U(p*, x) = y\}$$

Prefix complexity monotone in conditions

Definition (informal)

$x, y \in \mathcal{X}^*$, U is a universal twice-prefix machine, $p \in \{0,1\}^*$

$$KP_*(y|x*) = \min\{\ell(p) | U(p*, x*) = y\}$$

$$KP_*(y|xz*) \leq KP_*(y|x*)$$

$KP_*(\mu|x*)$ bound

Theorem

For any computable measure μ and any $x, y \in \mathcal{X}^*$

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \stackrel{+}{\leq} KP_*(\mu|x*) + KP(\lceil d_\mu(x) \rceil)$$

Corollary

$$\mathrm{Dist}(M, \mu|x_1 \dots x_n) \stackrel{+}{\leq} \min_{i \leq n} \{KP(\mu|x_1 \dots x_i) + KP(i) + KP(d_\mu(x_1 \dots x_i))\}$$

For μ-typical x, $\mathrm{Dist}(M, \mu|x) \leq KP(\mu|x') + O(\log \ell(x'))$

$KP_*(\mu|x*)$ bound

Theorem

For any computable measure μ and any $x, y \in \mathcal{X}^*$

$$\log_2 \frac{\mu(y|x)}{M(y|x)} \overset{+}{\leq} KP_*(\mu|x*) + KP(\lceil d_\mu(x) \rceil)$$

Corollary

$$\mathrm{Dist}(M, \mu|x_1 \dots x_n) \overset{+}{\leq} \min_{i \leq n}\{KP(\mu|x_1 \dots x_i) + KP(i) + KP(d_\mu(x_1 \dots x_i))\}$$

For μ-typical x, $\quad \mathrm{Dist}(M, \mu|x) \leq KP(\mu|x') + O(\log \ell(x'))$

Relations between Complexities

$U(p*, x*) = y$
$KP_*(y|x*)$

$U(p*, x) = y$
prefix

$U(p, x*) = y$

$U(p*, x*) = y*$

$U(p, x) = y$
plain

$U(p*, x) = y*$
monotone

$U(p, x*) = y*$

$U(p, x) = y*$
decision

Uspensky, Shen. Relations between Varieties of Kolmogorov Complexities. *Math. Systems Theory*, 1996

Relations between Complexities

$U(p*, x*) = y$
$KP_*(y|x*)$

$U(p*, x) = y$
prefix

$U(p, x*) = y$

$U(p*, x*) = y*$

$U(p, x) = y$
plain

$U(p*, x) = y*$
monotone

$U(p, x*) = y*$

$U(p, x) = y*$
decision

Uspensky, Shen. Relations between Varieties of Kolmogorov Complexities. *Math. Systems Theory*, 1996

Prefix complexity monotone in conditions

Definition (formal)

$E = \{\langle p, x, y\rangle\}$ is KP_*-correct if

1. $\langle p, x, y_1\rangle \in E,\ \langle p, x, y_2\rangle \in E \quad \Rightarrow \quad y_1 = y_2$;
2. $\langle p, x, y\rangle \in E \quad \Rightarrow \quad \langle p', x', y\rangle \in E \qquad \forall\, p' \sqsupseteq p,\, x' \sqsupseteq x$;
3. $\langle p, x', y\rangle \in E,\ \langle p', x, y\rangle \in E,\ \ p \sqsubseteq p',\ x \sqsubseteq x' \quad \Rightarrow$ $\langle p, x, y\rangle \in E$.

Let E be an optimal enumerable KP_*-correct set

$$KP_*(y|x*) = \min\{\ell(p) \mid \langle p, x, y\rangle \in E\}\,.$$