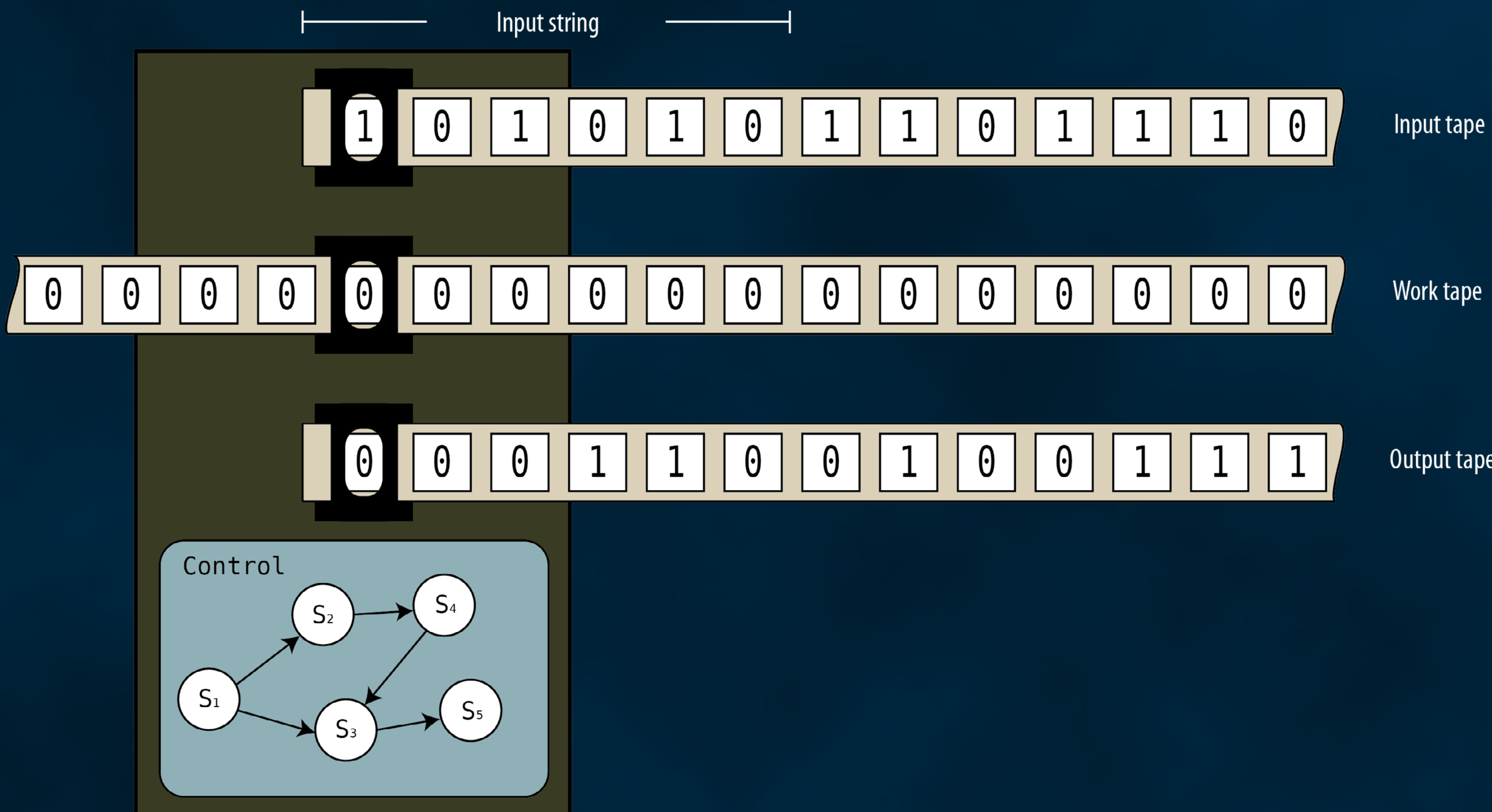
Minimum Message Length and Kolmogorov Complexity

C. S. Wallace and D. L. Dowe

Presentation by Alkis Gkotovos

Overview

Turing Machines



Two-part encoding

How do we acquire a hypothesis-based encoding of data in the Algorithmic Complexity framework?

Two-part encoding

How do we acquire a hypothesis-based encoding of data in the Algorithmic Complexity framework?

Idea: Use conditional Kolmogorov complexity

$$K_T(x \mid y) = \min\{l(p) \mid T(\langle y, p \rangle) = x\}$$

and interpret y as hypothesis and x as data.

Two-part encoding

How do we acquire a hypothesis-based encoding of data in the Algorithmic Complexity framework?

Idea: Use conditional Kolmogorov complexity

$$K_T(x \mid y) = \min\{l(p) \mid T(\langle y, p \rangle) = x\}$$

and interpret y as hypothesis and x as data.

Corresponding conditional algorithmic probability

$$P_T(x \mid y) = 2^{-K_T(x|y)}$$

Two-part encoding

How do we acquire a hypothesis-based encoding of data in the Algorithmic Complexity framework?

Idea: Use conditional Kolmogorov complexity

$$K_T(x \mid y) = \min\{l(p) \mid T(\langle y, p \rangle) = x\}$$

and interpret y as hypothesis and x as data.

Corresponding conditional algorithmic probability

$$P_T(x \mid y) = 2^{-K_T(x|y)}$$

Problem: Probability can never be 0, i.e. Popper-falsification not possible, because

$$K(x \mid y) < K(x) + O(1) \Rightarrow P_K(x \mid y) > P_K(x) + O(1)$$

Why?

Two-part encoding

How do we acquire a hypothesis-based encoding of data in the Algorithmic Complexity framework?

Idea: Use conditional Kolmogorov complexity

$$K_T(x \mid y) = \min\{l(p) \mid T(\langle y, p\rangle) = x\}$$

and interpret y as hypothesis and x as data.

Corresponding conditional algorithmic probability

$$P_T(x \mid y) = 2^{-K_T(x|y)}$$

Problem: Probability can never be 0, i.e. Popper-falsification not possible, because

$$K(x \mid y) < K(x) + O(1) \Rightarrow P_K(x \mid y) > P_K(x) + O(1)$$

Why? Hypothesis y acts as “extra info”, instead of assertively.

Two-part encoding

How do we acquire a hypothesis-based encoding of data in the Algorithmic Complexity framework?

Idea: Use conditional Kolmogorov complexity

$$K_T(x \mid y) = \min\{l(p) \mid T(\langle y, p\rangle) = x\}$$

and interpret y as hypothesis and x as data.

Corresponding conditional algorithmic probability

$$P_T(x \mid y) = 2^{-K_T(x|y)}$$

Problem: Probability can never be 0, i.e. Popper-falsification not possible, because

$$K(x \mid y) < K(x) + O(1) \Rightarrow P_K(x \mid y) > P_K(x) + O(1)$$

Why? Hypothesis y acts as “extra info”, instead of assertively.

Proposal: Have hypothesis be a prefix of input string p.
Force intended two-part encoding by imposing conditions on p.

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1) $T(p) = x$ p encodes x

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1) $T(p) = x$ — p encodes x

2) $l(p) < l(x)$ — some compression is achieved

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1) $T(p) = x$ — p encodes x

2) $l(p) < l(x)$ — some compression is achieved

3) $p = qr$ — two-part encoding

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1) $T(p) = x$ — p encodes x

2) $l(p) < l(x)$ — some compression is achieved

3) $p = qr$ — two-part encoding

4) $T(q) = \epsilon$ — hypothesis q is does not determine data

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1) $T(p) = x$ — p encodes x

2) $l(p) < l(x)$ — some compression is achieved

3) $p = qr$ — two-part encoding

4) $T(q) = \epsilon$ — hypothesis q is does not determine data

5) $T_q(rs) = xT_q(s)$ — reading r does not alter the state of T

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1)	$T(p) = x$	p encodes x
2)	$l(p) < l(x)$	some compression is achieved
3)	$p = qr$	two-part encoding
4)	$T(q) = \epsilon$	hypothesis q is does not determine data
5)	$T_q(rs) = xT_q(s)$	reading r does not alter the state of T
6)	$l(r) < K_T(x)$	hypothesis q is "significant"

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1)	$T(p) = x$	p encodes x
2)	$l(p) < l(x)$	some compression is achieved
3)	$p = qr$	two-part encoding
4)	$T(q) = \epsilon$	hypothesis q is does not determine data
5)	$T_q(rs) = xT_q(s)$	reading r does not alter the state of T
6)	$l(r) < K_T(x)$	hypothesis q is "significant"
7)	$x = x_1 \ldots x_n \Rightarrow \begin{cases} r = r_1 \ldots r_n \\ T_q(r_i) = x_i, \ i = 1 \ldots n \end{cases}$	conditionally independent sentences

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1)	$T(p) = x$	p encodes x
2)	$l(p) < l(x)$	some compression is achieved
3)	$p = qr$	two-part encoding
4)	$T(q) = \epsilon$	hypothesis q is does not determine data
5)	$T_q(rs) = xT_q(s)$	reading r does not alter the state of T
6)	$l(r) < K_T(x)$	hypothesis q is "significant"
7)	$x = x_1 \dots x_n \Rightarrow \begin{cases} r = r_1 \dots r_n \\ T_q(r_i) = x_i, \; i = 1 \dots n \end{cases}$	conditionally independent sentences
8)	$\begin{matrix} x' = x^{(1)}x^{(2)} \\ j' = j^{(1)}j^{(2)} \end{matrix} \Rightarrow \begin{matrix} T_q(j^{(1)}) = x^{(1)}, \; j^{(1)} < K_T(x^{(1)}) \\ T_q(j^{(2)}) = x^{(2)}, \; j^{(2)} < K_T(x^{(2)}) \end{matrix}$	hypothesis q is "general"

Two-part encoding

Input p is an acceptable MML message encoding data string x, if

1) $T(p) = x$ — p encodes x

2) $l(p) < l(x)$ — some compression is achieved

3) $p = qr$ — two-part encoding

4) $T(q) = \epsilon$ — hypothesis q is does not determine data

5) $T_q(rs) = xT_q(s)$ — reading r does not alter the state of T

6) $l(r) < K_T(x)$ — hypothesis q is "significant"

7) $x = x_1 \dots x_n \Rightarrow \begin{cases} r = r_1 \dots r_n \\ T_q(r_i) = x_i, \; i = 1 \dots n \end{cases}$ — conditionally independent sentences

8) $\begin{matrix} x' = x^{(1)}x^{(2)} \\ j' = j^{(1)}j^{(2)} \end{matrix} \Rightarrow \begin{matrix} T_q(j^{(1)}) = x^{(1)}, \; j^{(1)} < K_T(x^{(1)}) \\ T_q(j^{(2)}) = x^{(2)}, \; j^{(2)} < K_T(x^{(2)}) \end{matrix}$ — hypothesis q is "general"

9) No prefix of q satisfies all the above conditions — all of q is required