Kolmogorov-Chaitin Complexity

CPSC 501

Grigorii Guz, 07.12.2021

Information

- Different definitions, depending on:
  - The *kind* of entity whose information content we are trying to measure.
  - What are we trying to quantify the information content *for*?
- We could measure the amount of information coming from a “communication channel” (Entropy):
  - Assume it produces 2 messages: 0110, which means “hello”, and 1110, which means “goodbye”. Then the amount of information that this channel can produce is 1 bit.

Information of an Object

- A quantity of information contained in an object is the size of the object's *smallest description*.
  - Description is a precise and unambiguous characterization of the object, from which it can be recreated.
    - "01010101010101010101"
      - = "01" * 10
    - "1101010001"
      - = "1101010001"
    - Your DNA: 3 million nucleotides (ATCG letters) vs. 37.2 trillion cells.

Characterization of Description Procedure

- One way to produce a description of a binary string x is by a corresponding Turing machine M , producing x on blank input.
  - The encoding $\langle M \rangle$ itself is a description of x.
  - Wasteful: for a string of size n, we would need n states and as many rows in the transition table.
  - Example: a long string of 1's.

A Shorter Description Procedure

- Characterization of a string x is a Turing machine M and a binary input w to M.
    - The length of the description is the length of the string $\langle M, w \rangle$, which can be further simplified to $\langle M \rangle w$.
    - Let x be a string of n consecutive 1s, M be a TM that copies the input, and w be a string of $\frac{n}{2}$ consecutive 1s. Then $\langle M, w \rangle = \frac{n}{2} + c < n$ for n large enough.
- The minimal description $d(x)$ of a string x is then the shortest string $\langle M \rangle w$.

Minimal Description Length

- The ***minimal description*** $d(x)$ of a string x is then the shortest string $\langle M \rangle w$ such that the TM M halts on input w with x on the tape.
- The ***descriptive complexity (Kolmogorov complexity)*** of x, written $K(x)$, is the length of the minimal description $|d(x)|$.

Properties of KC [1]

- Theorem 1. $\exists c \in \mathbb{N}, \forall x \in \{0,1\}^*, K(x) \leq |x| + c$
  - Consider a TM M that halts immediately after it starts. Then the tape remains unchanged. Let $\langle M \rangle = c$.
- Theorem 2. $\exists c \in \mathbb{N}, \forall x \in \{0,1\}^*, K(xx) \leq K(x) + c$
  - Consider $d(x) = \langle N, w \rangle$ as a minimal description of x and a Turing machine M that
    - 1. Runs N on w, which produces x.
    - 2. Doubles the input on the tape to be xx.
  - This gives the description of xx as $\langle M \rangle d(x)$, and $\langle M \rangle d(x) = |\langle M \rangle| + |d(x)| = c + K(x)$.

Properties of KC [2]

- Theorem 3. $\exists c \in \mathbb{N}, \forall x, y \in \{0,1\}^*, K(xy) \leq 2K(x) + K(y) + c$
  - Consider a TM M breaking down its input w into 2 parts. The first part of w is a doubled $d(x)$, which is terminated by 01 and then followed by $d(y)$. Then run M to obtain xy. Since $d(x)$ is doubled, we have:
    - $|\langle M, w\rangle| = |\langle M\rangle| + 2d(x) + d(y) = c + 2K(x) + K(y)$

Optimality

- Let $p : \{0,1\}^* \rightarrow \{0,1\}^*$ be a computable function. A minimal description of x with respect to p, $d_p(x)$, is the smallest string s s.t. $p(s) = x$. Let $K_p(x) = |d_p(x)|$.

- Theorem 4. For any description language p there exists a constant c depending only on p such that

  - $\forall x, K(x) \leq K_p(x) + c$

  - Take any description language p and consider the TM M, which given an input w, prints $p(w)$. Then $|d(x)| = |\langle M \rangle| + |d_s(x)|$.

Incompressibility of Some Strings

- Let x be a string, and a constant c. Then x is c*-compressible* if and only if
  - $K(x) \leq |x| - c$
- If it is not c-compressible, it is said to be c*-incompressible.* Further, if it is incompressible by 1, it is called *incompressible.*

Incompressibility of Some Strings

- Theorem. The proportion of c-compressible strings of length n is at most $\frac{1}{2^c}$.
  - $\frac{|\{x : K(x) \leq |x| - c, |x| = n\}|}{|\{x : |x| = n\}|}$
  - $= \frac{|\{x : K(x) \leq |x| - c, |x| = n\}|}{2^n}$
  - $\leq \frac{2^{n-c}}{2^n} = \frac{1}{2^c}$
- The proportion of strings 10-compressible length-20 strings is at most $\frac{1}{1024}$.

Computability of Kolmogorov complexity

- Theorem: $K(x)$ is uncomputable.
  - Assume there is a TM p s.t. $p(x) = K(x)$ for any binary string x. Pick $d \in \mathbb{N}$ such that $d > \log d + |p| + c$.
  - Let q be a program that generates the first binary string y s.t. $K(y) = d$: just iterate over all binary strings in lexicographic order, and pick the first one of required complexity (by applying p).
  - What is $|q|$? It is $|p| + \log d + c$, where $\log d$ is the number of bits to write the number d, and c is the length of the rest of the program.
  - Clearly, q is a description of y, so $d = K(y) \leq |q| = |p| + \log d + c$.
  - But also, $d = K(y) > \log d + |p| + c$, which is a contradiction!

The Halting Problem

- Theorem: If $HALT_{TM} = \{\langle M, w\rangle \mid M \text{ halts on } w\}$ is decidable, then $K(x)$ is computable.
  - Assume there is a TM N that decides $HALT_{TM}$.
  - Now, iterate over all possible $\langle M, w\rangle$ in lexicographic order.
  - If N does not halt on those, continue. If it halts, check if M produces x on input w.

References

- Peter Grunwald and Paul Vitanyi. *Shannon Information and Kolmogorov Complexity*. https://homepages.cwi.nl/~paulv/papers/info.pdf
- Michael Sipser. *Introduction to the Theory of Computation.* 3rd ed.
- Shai Ben-David. Lecture 22 of *Logic of Computation*, CS245, Fall 2015, University of Waterloo.