STOC LECTURE 26

Note Title 5/18/2006

Today

- Computational Perspectives (cntd.)
  - Kolmogorov Complexity
  - Pseudorandomness
  - Information vs. Knowledge
- Summary.

Motivating question:

How would you "compress" the string

0100 01 10 11 000 001010 011101110111...

n bits.

How well can you compress the bible?

"Shannon Coding": Single string?

Needs 0 bits!

Single string from larger space of possibilities ... Given me list & I'll tell you ...

"Universal Coding" a la Lempel-Ziv:

- ✓ Can talk about compression
- ✓ But can't compress this string much
- ✓ No real repetitions.

But actual compression:

"Enumerate binary strings in lex. order" for "n" bits.

Compression length = $\log_2 n$ + $O(1)$ bits

Is this "legitimate"

What properties does this have.

Investigated by Solomonoff - Kolmogorov - Chaitin in 1960's.

Defn: Represent string x by computer program that outputs x and stops. $K(x) = |\text{prog}|$

Is this well-defined?

Depends on programming language!

E.g. in Faith based - C++ "print bible"
print entire bible. $\Rightarrow$ Com (Bible) = $O(1)$

So programming languages can make difference

for language L $K_L(x)$ seems well-defined

But how does $K_L(x)$ compare with

$K_{L'}(x)$?

Informal theorem: For "complete"

("universal") languages L, $K_L(x)$ is

nearly same ....

Formally $\exists$ (unreal) language U

s.t $\forall$ language L,

$\exists$ $C_{L,U}$

$\forall$ x

$$K_U(x) \leq K_L(x) + C$$

Importance of order

① $\exists U \; \forall L, x, \; \exists C$ would be trivial since

$K_U(x) = |x|$ would satisfy this.

② $\exists U, C \; \forall L, x$ ---- not true.

(Faith based C++) ...

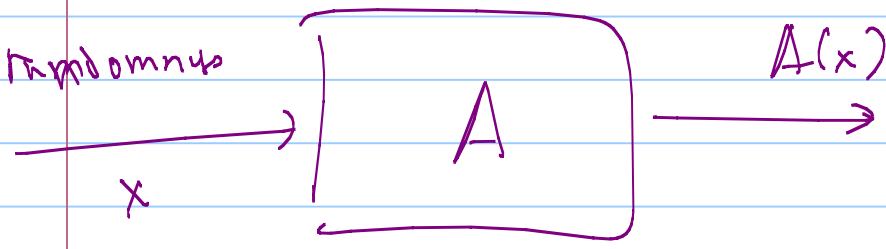
⇓

(x-based C++) ....

etc ...

Ex: Kolmogorov Compression ≤ Lempel-Ziv ....

Kolmogorov Probability measures



(can assume iid bits)

How would you compress $A(x)$, given the description of this process?

Good compression: $\langle n, x_n \rangle$

n = # bits of x read by A

x_n = prefix of length n.

(optimal if A invertible.)

Comparing for dist sampled by A.

But then Kolmogorov complexity is optimal to within constant ....

$\Rightarrow$ Notion of universal probability ...

$$P_U(x) = \sum_{\substack{\text{programs } p \\ \text{s.t.} \\ U(p) = x}} 2^{-l(p)}$$

$\forall L \; \exists C \; \forall x \; \ldots$

$$P_L(x) \leq C_L \cdot P_U(x)$$

K - Complexity Summary

- Best goal for "Compression"

- Unachievable .... ∃ short programs whose output is undefined....
  E.g. "Shortest word whose Kolmogorov complexity is greater than thousand bits".

LZ	vs.	K
feasible goal		infeasible goal
weak comp.		strong comp.

Why not ask for "best feasible compression"?

E.g. "Shortest program that produces X in $|X|^2$ steps" ...

Can be "computed" in time $\approx 2^{n^2}$ steps

"Resource-Bounded Kolmogorov Complexity".

---

feasible computation vs. information

Eg: Suppose $N = p \cdot q$

is "N" = "p" · "q"

↑ in binary

Classical notion of information:

If A is invertible

Information in x

$=$

Info in $A(x)$

But not wrt feasible computation.

Notion of feasible computation changes

- Amount of information in string
- Seeming amount of randomness ...

Cryptography

Knowledge $\leq$ Information $\Rightarrow$ Comm.

$+$ $+$

Pseudorandomness $\geq$ randomness

$=$ $=$

n n

Summary of Course

- Entropy
- Mutual Information
- AEP : for many nice sources of randomness .... distribution dominated by near uniform dist on some (small) subset.
- Source Coding ; Universal Coding ;
- Channel Coding.
- Continuous variables + Differential Entropy
- Network Inf. Theory.

Main Elements we missed

- Rate Distortion Theory.
- Applications . .

Hope : if you look at situations involving signals ⇒ think prob.
in prob. ⇒ think info
↙ entropy

Future Courses :

Comm : 6.451, 6.452 . . . .

Computation : 6.046 algorithms
. 6.840 complexity
. 6.841 Adv. complexity