Lens at other fields

A brief introduction to
Descriptive Complexity
and the Probabilistic Method

Objective

- We have completed the first part of our course--Algorithm Design Techniques, and are now being at a point to start on the way to NP-completeness.
- At this point, we'd better to have a relaxation from the intensive study, so as to
  - celebrate for having swallowed the first part
    - Though we need time to digest
  - accumulate energy for our future journey.
- Thus, this lecture intends to provide a leisurely, informal peek at other two active and algorithm-related fields, in order to
  - amuse our mind,
  - broaden our sight.

Descriptive complexity

- Equally appropriate titles:
  - "Kolmogorov Complexity",
  - "Algorithmic Information Theory",
  - "Algorithmic Complexity",
  - "Program-Size Complexity".
  - …
- Each name represents
  - a variation of the basic idea, or
  - a different point of departure.
- Main contributor: A.N. Kolmogorov (1903-1987).
- Current most active researcher: Ming Li.

Two binary strings

- What is the difference between the next two?
  1. 010101010101010101010101010101010101010101
  2. 101011100100110011111011011111000010100111
- Possible answers:
  1. The first is regular, the second is random.
  2. The first is easy to member, the second not.
  3. The first can be described succinctly, the second has to be stated by spelling out the whole string.
  4. The first can be compressed, the second not likely.
  5. The first contains few information, the 2nd much.

Description methods and object complexity

- For a set X of objects, a specification method D for X gives each object $x \in X$ at least one description y—denoted D(y)=x.
  - D can be
    - A natural language such as English, Chinese,…
    - A computer language, a description y of x is the program that produces x.
    - …
- The length of the shortest description of object x is called the descriptive complexity of x, denoted by $C(x)=\min_{D(y)=x}|y|$.

Is C(x) well-defined?

- In some sense, Yes.
- The shortest description length of an object is an intrinsic attribute of the object.
  - Independent of the particular description method,
  - Any two reasonable methods D1,D2 give the complexity of a same object within additive constant. That is, $\exists c$, $C_{D1}(x) \le C_{D2}(x)+c$
- But, beware of falling into a disturbing trap--Richard-Berry paradox, it defines a natural number as
  - The least natural number that cannot be described in less than 78 characters.

Incompressibility

- For every n, there is a string x of length n such that C(x)≥|x|.
  - Proof: by counting
- This yields a simple but powerful proof technique—the incompressibility method, a general purpose tool, comparable to the pigeon-hole principle

Simple application

——number theory

- For infinitely many natural number n, the number of primes $\leq n$ is at least $\log n/\log\log n$.
  - Let n be incompressible, i.e., n cannot be described in $< \log n$ bits
  - Assume that $p_1, p_2, \ldots p_m$ are all the primes $<n$.
  - Then, $n = p_1^{e1}, p_2^{e2}, \ldots p_m^{em}$.
  - We can describe n by $(e1, e2, \ldots, em)$
  - Each $ei \leq \log n$, can be represented by $\log\log n$ bits.
  - The description of n is given in $m\log\log n$ bits.
  - $m\log\log n \geq \log n$, giving us that $m \geq \log n / \log\log n$.
- A slightly more complicated encoding can improve the above result to $n/\log^2 n$.

Simple application

——compact routing

- There is an n-node network such that any all-shortest-path routing function must consume at least $(n-1)/2$ bits at some node.
  - The topological structure of network G=(V, E) can be recovered by combination of all the routing functions.
  - The is total $2^{\binom{n}{2}} = 2^{\frac{n(n-1)}{2}}$ networks of n nodes.
  - Some G must use at least $\log 2^{\binom{n}{2}} = \frac{n(n-1)}{2}$ bits to be described.
  - For such G, some node must has a space of (n-1)/2 bits for its routing function.

Fruits

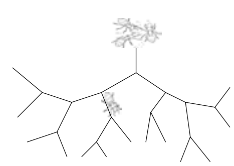
- Fruitful in
  - Probabilistic theory
    - The philosophical notion of randomness
  - Information theory
  - Computer science
  - Physics
  - Biologics
  - ….

Compression in Nature

- Learning, in general, appears to
  - Involve compressions of observed data or the results of experiments.
  - If the learner cannot compress the data, s/he does not learn.
- We often compress information that is presented to us by the environment.
  - E.g., Science may be regarded as the art of data compression
    - Compress a great number of experimental data into a short natural law.
- Perhaps animals do this as well
  - DNA
  - Ant

Compression by ants

- An experiment shows that ants are able to compress information
- The experiment was reported by Zh.I. Reznikova and B.Ya. Ryabko at Problems of Information Transmission 22:3, 1986,245-249.



Maze: a binary tree constructed with matches, floating on water, Connected to the nest

The probabilistic method

- A tool that is
  - powerful and widely used,
  - in recent years developed rapidly.
    - Reason: the important role of randomness in CS.
- This method was initiated by Paul Erdos.
  - Erdos method?
- Basic idea:
  - in order to prove the existence of a structure with certain property, we
    - define a property space and
    - show that a random chosen structure has the desired property with positive probability.

Example— Ramsey number

- Prove $R(k,k)>\lfloor 2^{k/2}\rfloor$ if $k>3$. (The Ramsey number $R(k,l)$ is define to be the smallest n such that in K_n for any two-coloring of the edges by red and blue, either there is a red K_k or there is a blue K_l)
- Proof: Consider a random two-coloring of K_n obtained by coloring each edge independently either red or blue, each color is equally likely.
  - For any set R of k vertices, let A_R be the event that the induced subgraph on R is monochromatic (all its edges are colored same).
  - Clearly, $P(A_R)=2^{1-\binom{k}{2}}$.
  - Since there are $\binom{n}{k}$ possible choices of A_R, the probability of at least one of A_R occurs is at most $\binom{n}{k}2^{1-\binom{k}{2}}$.
  - If $k>3$ and $n=\lfloor 2^{k/2}\rfloor$ then we will have $\binom{n}{k}2^{1-\binom{k}{2}}<\frac{2^{1+\frac{k}{2}}}{k!}\frac{n^k}{2^{k^2/2}}<1$,
    - implying that with a positive probability, no event A_R occurs.
  - So it is must that $R(k,k)>n=\lfloor 2^{k/2}\rfloor$

Example— large cuts

- There exists a cut C for G=(V,E) with |C|>|E|/2. (a cut is a set of the edges that connect vertices of U with the vertices of V-U)
- Proof: randomly and independently include each vertex u in to U with probability 1/2. Let C={(x,y)|exactly one of x, y in U}.
  - We need only to prove that P(|C|>|E|/2)>0.
    - For e={x,y}∈E, define random variable
    - $X_e=\begin{cases}0 & e\in C\\ 1 & e\notin C\end{cases}$ and let $X=\sum_{e\in E}X_e$.
    - Clearly, X=|C| and the expectation $E(X_e)$=0*1/2+1*1/2=1/2,
    - and $E(|C|)=E(X)=E\left(\sum_{e\in E}X_e\right)=\sum_{e\in E}E(X_e)=\sum_{e\in E}\frac{1}{2}=\frac{|E|}{2}$,
    - which means that P(|C|>|E|/2)>0

Remark

- The probabilistic method is to prove the existence of an object in a nonconstructive way.
- The probabilistic method is extremely useful in Combinatorics, Graph Theory, Number Theory, Geometry, etc.
- More recently, it has been applied in
  - the development of efficient algorithms, and
  - in the study of various computational problems.