Midterm 2 Study Guide, by Ivan "Vania" Smirnov

Error Correcting Codes:

Erasure Errors

- use polynomial interpolation to get missing point.
- Sending n packets, guard against k errors, send $n+k$ packets.
- GF(q) - each packet can be encoded mod q, so q > largest number in data
- send packets,ensure $n+k \leq q$.
- use delta reconstruction $\Delta_3(x) = \frac{(x-a_1)(x-a_2)(x-a_4)}{(a_3-a_1)(a_3-a_2)(a_3-a_4)}$
- Then add all them all up: $y_1\Delta_1 + y_2\Delta_2 ...$ to get original polynomial.

General Errors

- n length message, k errors, send $n+2k$ message
- $Q(x) = P(x)E(x) \rightarrow Q(x)/E(x) = P(x)$
- Sending the message:
  1. get n points, fine deg(n-1) polynomial.
  2. evaluate $2k$ more points.
  3. send $P(i)$ for $i \in 0, 1, ...(n+2k)$
- Decoding the message:
  1. Note - rememeber $GF(p)$, so mod stuff!
  2. get $n+2k$ points
  3. $E(x)$ of degree k, $(E(x) = x^k + b_{k-1}x^{k-1} + ...b_1x + b_0)$
  4. $Q(x)$ is degree $n+k-1$, $(Q(x) = a_{n+k-1}x^{n+k-1} + ...a_1x + a_0)$
  5. for each point x_i, subsitute in to $Q(x) = r_xE(x)$, (r_x is x_i)
  6. Solve system for $a_1, a_2...b_1, b_2...$
  7. These are coefficents of $Q(x)$ and $E(x)$. $\frac{Q(x)}{E(x)} = P(x)$

Graphs

Notation: G(V,E) - Graph, Vertexes, Edges. E is subset of V x V (Cartesian product) Can be bidirectional or directed.

Terms used:

1. path - sequence of edges
2. simple path - no repeated vertices
3. cycle/circuit - path that begins and ends on same vertex
4. connected - some path between any two vertices
5. incident - edge "incident" to v_1 and v_2 if it connectes them
6. Eulerian path - path that uses each edge exactly once
7. Eulerian tour - cycle that uses each edge exactly once
8. Hamilton path - visits each vertex exactly once
9. Hamiltonian cycle - cycle, same as above.
10. XX - as you can see, hamiltonian = vertices, euclidean = edges.

Properties

1. Eulerian tour only exists if undirected graph connected and even degree.

Hypercubes

1. n-bit strings, each connected to points differing from it by one bit, so each vertex degree n. Very efficient, little bottlenecks. (010 $\rightarrow$ 011)
2. two copies of n-1 matrix, connected by 0x-1x
3. total number of edges is $n2^{n-1}$
4. properties - submatrixes, etc.

Probability Basics

1. Coin flipping - normal curve, std dev is $\sqrt{n}$
2. Given a confidence (98%), set σ to $1 - .98 = 0.02$, and a desired error ε, then $n = \frac{1}{4\varepsilon^2\sigma}$

Counting

1. Sampling w/o replacement, order matters - n permutation $k = \frac{n!}{(n-k)!}$
2. Sampling w/o replacement, order doesn't matter - n choose k, $\frac{n!}{(n-k)!k!}$
3. Sampling with replacement: order matters. n items, k trials, n^k possibilities.
4. Sampling with replacement, order doesn't matter: n bins, k items. $\binom{n+k-1}{k}$

Formulas

- Bayesian Inference: $\Pr[A|B] = \frac{Pr[B|A]Pr[A]}{Pr[B|A]Pr[A]+Pr[B|!A](1-Pr[A])}$ (denominator is Pr[B])
- $\Pr[A|B] = \frac{P[B|A]P[A]}{P[B]}$
- Conditional probability - A given B: $\Pr[A|B] = \frac{P[A\cap B]}{\Pr[B]}$
- Total Probability rule: $\Pr[A] = \Pr[A|B]\Pr[B] + \Pr[A|\overline{B}]\Pr[\overline{B}]$
- $\Pr[A \cap B] = \Pr[A] * \Pr[B]$, intersection, AND. (assume independent)
- $\Pr[A \cup B] = \Pr[A] + \Pr[B]$, union, OR (independent)
- $\Pr[A \cap B] = \Pr[A]\Pr[B|A]$, intersection(dependent)
- $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cup B]$, union(dependent)
- event C we get exactly r results of probability p given n trials = $P[C] = \binom{n}{r}p^r(1-p)^{n-r}$
- For events $A_1, \ldots, A_n$ in some probability space, we have $Pr[\cup_{i=1}^{n} Pr[A_i] = \sum Pr[A_i] - \sum Pr[A_i \cap A_j] + \sum Pr[A_i \cap A_j \cap A_k] - \cdots \pm \cup_{i=1}^{n} Pr[A_i]$. (basically count all individual events, subtract intersections of pairs, add back intersectons of triples, repeat alternating.)

Terms

- Probability space - sample space (all possible results) coupled with probability of each
- event - set of outcomes (more than 2 heads, exactly 1 tail, etc.)
- outcome - specific result (THT)
- probability = $\frac{desired}{allpossible}$

Example:

Given P[Rain] = 0.6, and Charlie predicts rain with correctness .7, and false positive .4, and Kevin predicts rain with correctness .6, and false positive .1, find who's reliable and P[R|C].

- Total Probability: P[C] = P[C|A]P[A] + P[C|$\overline{A}$]P[$\overline{A}$] (same for P[K])
- If C says Rain, what's the chance of it being true?

Balls and Bins

- Pr[bin 1 is empty] = $(\frac{n-1}{n})^m = (1 - \frac{1}{n})^m$
- Pr[first k out of n bins empty] = $(1 - \frac{k}{n})^m$
- Given k out of n bins empty, Pr[(k+1)th bin empty] = $\frac{(1-\frac{k+1}{n})^n}{(1-\frac{k}{n})^n} = (\frac{m-k-1}{m-k})^n$
- Birthday paradox. Probability NOT same birthday is: $\frac{365*364*\ldots*(365-n+1)}{365^n}$, so with 1 - *this* we get 50% with 23 people.

1. Relax.
2. You will do GREAT!
3. The "A" is yours.