# Reliable Meaningful Communication

**Madhu Sudan**
Microsoft Research

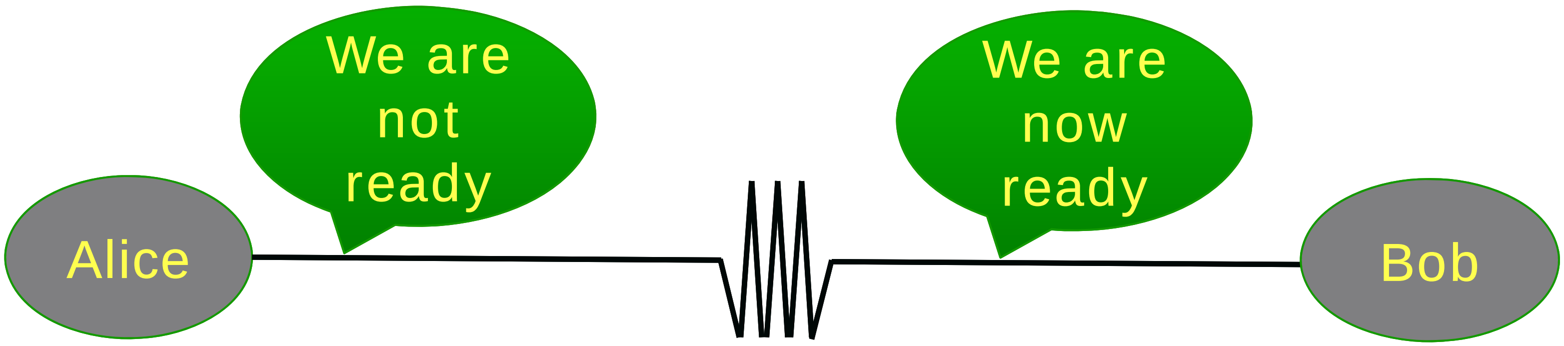
# This Talk

- Part I: Reliable Communication
  - Problem and History (briefly)
- Part II: Recovering when errors overwhelm
  - Sample of my work in the area
- Part III: Modern challenges
  - Communicating amid uncertainty

# Part I: Reliable Communication

# Reliable Communication?

- Problem from the 1940s: Advent of digital age.




- Communication media are always noisy
  - But digital information less tolerant to noise!

# Reliability by Repetition

- Can repeat (every letter of) message to improve reliability:

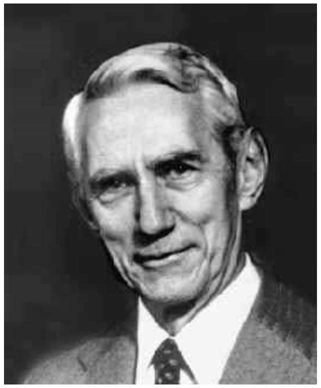
WWW EEE AAA RRR EEE NNN OOO WWW ...

↓

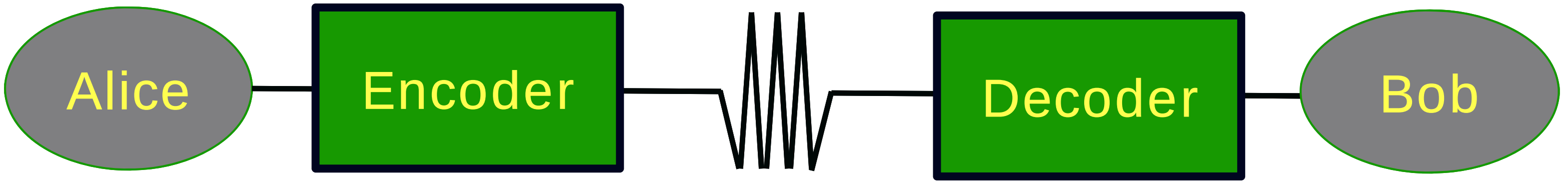
WXW EEA ARA SSR EEE NMN OOP WWW ...

- Elementary Calculations:
  - ↑ repetitions ⇒ ↓ Prob. decoding error; but still +ve
  - ↑ length of transmission ⇒ ↑ expected # errors.
  - Combining above: Rate of repetition coding $\rightarrow 0$ as length of transmission increases.
- Belief (pre1940):
  - Rate of any scheme $\rightarrow 0$ as length $\rightarrow \infty$

# Shannon's Theory [1948]



- Sender "Encodes" before transmitting
- Receiver "Decodes" after receiving




- Encoder/Decoder arbitrary functions.

$$E: \{0,1\}^k \to \{0,1\}^n$$
$$D: \{0,1\}^n \to \{0,1\}^k$$

  - Rate $= \frac{k}{n}$;
  - Requirement: $m = D(E(m) + \mathrm{error})$ w. high prob.
- What are the best $E, D$ (with highest Rate)?

## Shannon's Theorem

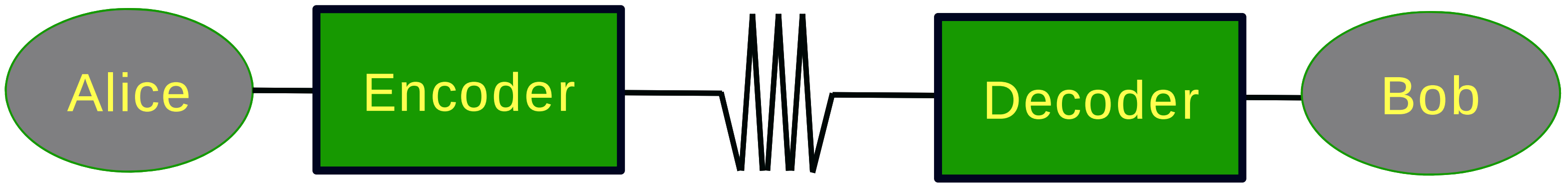
- If every bit is flipped with probability $p$
  - Rate $\rightarrow 1 - H(p)$ can be achieved.

$$H(p) \triangleq p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$$

  - This is best possible.
  - Examples:
    - $p = 0 \Rightarrow Rate = 1$
    - $p = \frac{1}{2} \Rightarrow Rate = 0$
    - Monotone decreasing for $p \in (0, \frac{1}{2})$
    - Positive rate for $p = 0.4999$ ; even if $k \rightarrow \infty$

# Shannon's contributions

- Far-reaching architecture:




- Profound analysis:
  - First (?) use of probabilistic method.
- Deep Mathematical Discoveries:
  - Entropy, Information, Bit?

# Challenges post-Shannon

- Encoding/Decoding functions not “constructive”.
  - Shannon picked $E$ at random, $D$ brute force.
  - Consequence:
    - $D$ takes time $\sim 2^k$ to compute (on a computer).
    - $E$ takes time $2^{2^k}$ to find!
- Algorithmic challenge:
  - Find $E, D$ more explicitly.
  - Both should take time $\sim k, k^2, k^3$ … to compute

## Progress 1950-2010

- Profound contributions to the theory:
  - New coding schemes, decoding algorithms, analysis techniques …
  - Major fields of research:
    - Communication theory, Coding Theory, Information Theory.
- Sustained Digital Revolution:
  - Widespread conversion of everything to “bits”
  - Every storage and communication technology relies/builds on the theory.
  - “Marriage made in heaven” [Jim Massey]

# Part II: Overwhelming #errors

# Explicit Codes: Reed-Solomon Code

- Messages = Coefficients of Polynomials.
  - Example:
    - Message = $(100,23,45,76)$
    - Think of polynomial $p(x) = 100 + 23x + 45x^2 + 76x^3$
    - Encoding: $(p(1), p(2), p(3), p(4), \dots, p(n))$
    - First four values suffice, rest is redundancy!
  - (Easy) Facts:
    - Any $k$ values suffice where $k$ = length of message.
    - Can handle $n - k$ erasures or $(n-k)/2$ errors.
    - Explicit encoding = polynomial evaluation ✓
    - Efficient decoding? [Peterson 1960]

# Overwhelming Errors? List Decoding

- Can we deal with more than 50% errors?
  - $\frac{n}{2}$ is clearly a limit – right?
    - First half = evaluations of $p_1$
    - Second half = evaluations of $p_2$
    - What is the right message: $p_1$ or $p_2$?

- $\frac{n}{2}$ (even $\frac{n-k}{2}$ ) is the limit for "unique" answer.
- List-decoding: Generalized notion of decoding.
  - Report (small) list of possible messages.
  - Decoding "successful" if list contains the message polynomial.

# Reed-Solomon List-Decoding Problem

- Given:
  - Parameters: $n, k, t$
  - Points: $(x_1, y_1), \ldots, (x_n, y_n)$ in the plane

    (finite field actually)
- Find:
  - All degree $k$ poly's that pass thru $t$ of $n$ points
    - i.e., all $p$ s.t.
      - $\deg(p) < k$
      - $\#\{\, i \mid p(x_i) = y_i \,\} \geq t$

# Decoding by example + picture [S'96]

$n = 14; k = 1; t = 5$

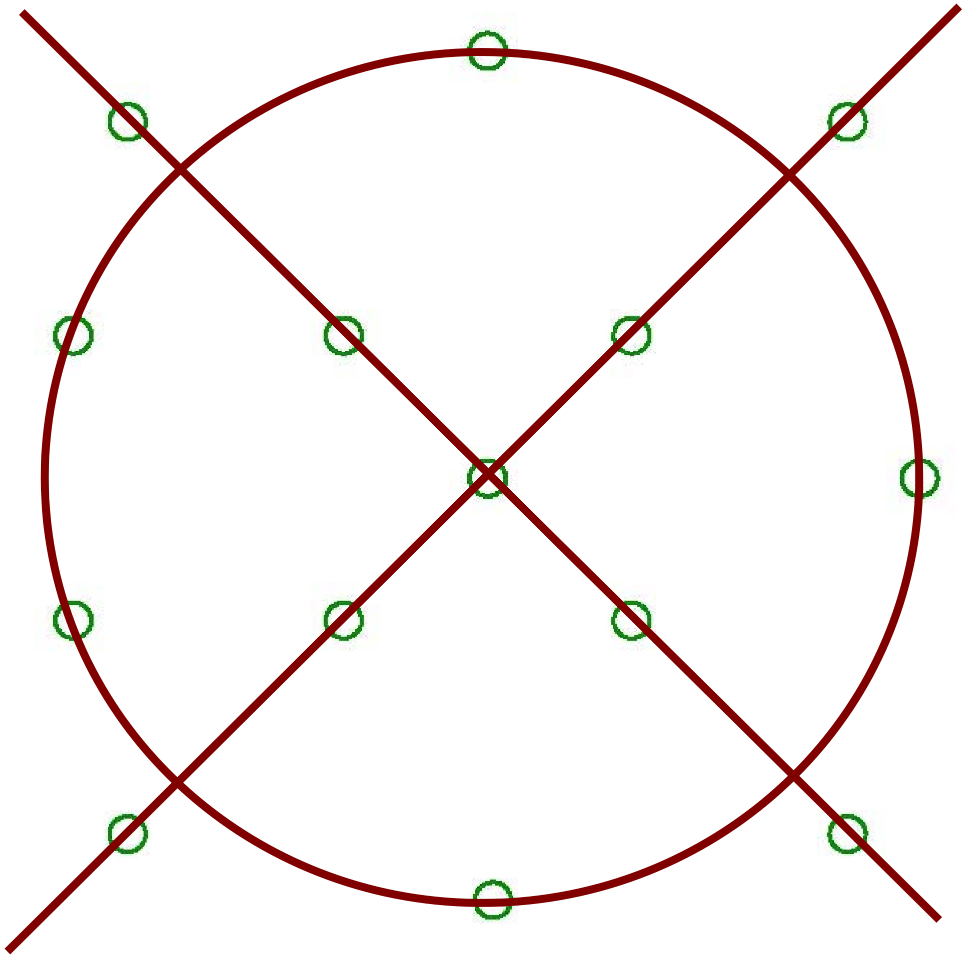
Algorithm idea:

- Find algebraic explanation of all points.

$$x^4 - y^4 - x^2 + y^2 = 0$$

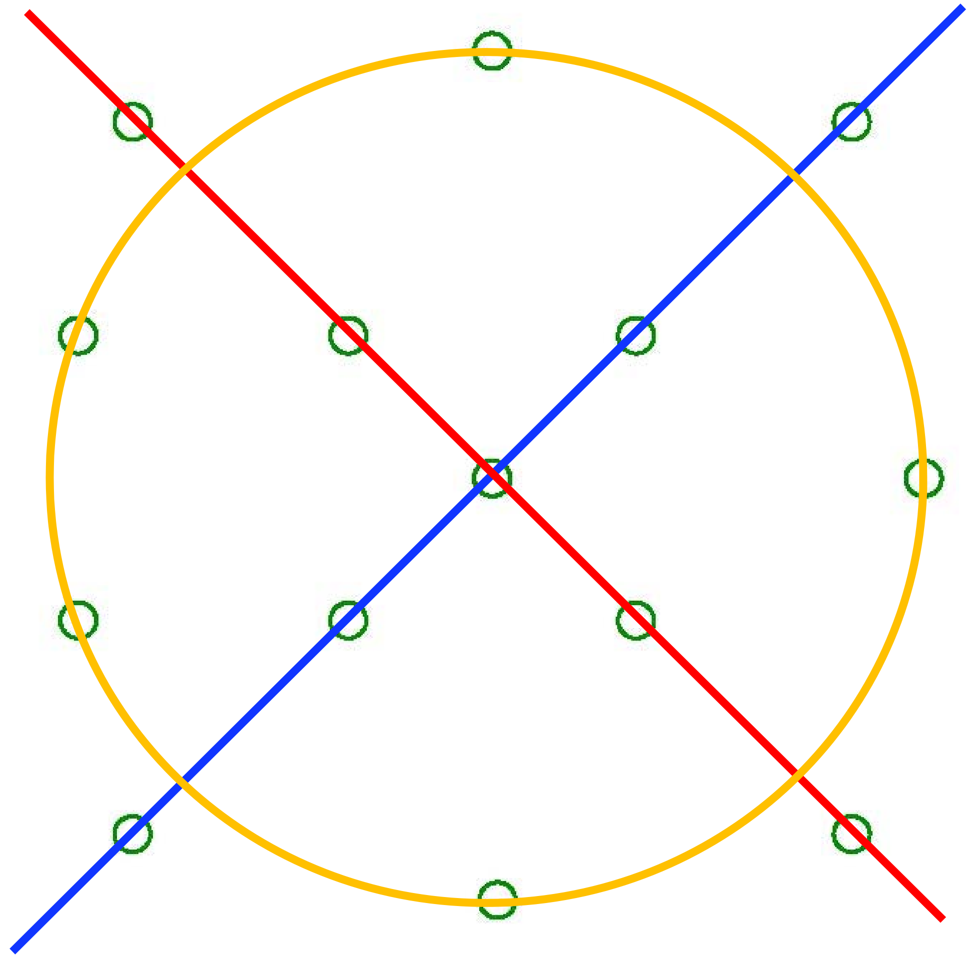
- Stare at the solution ☺ (factor the polynomial)

$$(x+y)(x-y)(x^2+y^2-1)$$



# Decoding by example + picture [S'96]

$n = 14; k = 1; t = 5$



Algorithm idea:

- Find algebraic explanation of all points.

$$x^4 - y^4 - x^2 + y^2 = 0$$

- Stare at the solution ☺ (factor the polynomial)

$$(x + y)\,(x - y)\,(x^2 + y^2 - 1)$$

## Decoding Algorithm

- Fact: There is always a degree $2\sqrt{n}$ polynomial thru $n$ points
  - Can be found in polynomial time (solving linear system).
- [80s]: Polynomials can be factored in polynomial time [Grigoriev, Kaltofen, Lenstra]
- Leads to (simple, efficient) list-decoding correcting $\kappa$ fraction errors for $\kappa \to 1$

# Part III: Modern Challenges

## Communication Amid Uncertainty?

# New Kind of Uncertainty

- Uncertainty always has been a central problem:
  - But usually focusses on uncertainty introduced by the channel
  - Rest of the talk: Uncertainty at the endpoints (Alice/Bob)
- Modern complication:
  - Alice+Bob communicating using computers
  - Huge diversity of computers/computing environments
  - Computers as diverse as humans; likely to misinterpret communication.
- Alice: How should I "explain" to Bob?
- Bob: What did Alice mean to say?

# New Era, New Challenges:

- Interacting entities not jointly designed.
  - Can't design encoder+decoder jointly.
  - Can they be build independently?
  - Can we have a theory about such?
    - Where we prove that they will work?

    - Hopefully:
      - YES
      - And the world of practice will adopt principles.

## Example Problem

- Archiving data
  - Physical libraries have survived for 100s of years.
  - Digital books have survived for five years.
  - Can we be sure they will survive for the next five hundred?

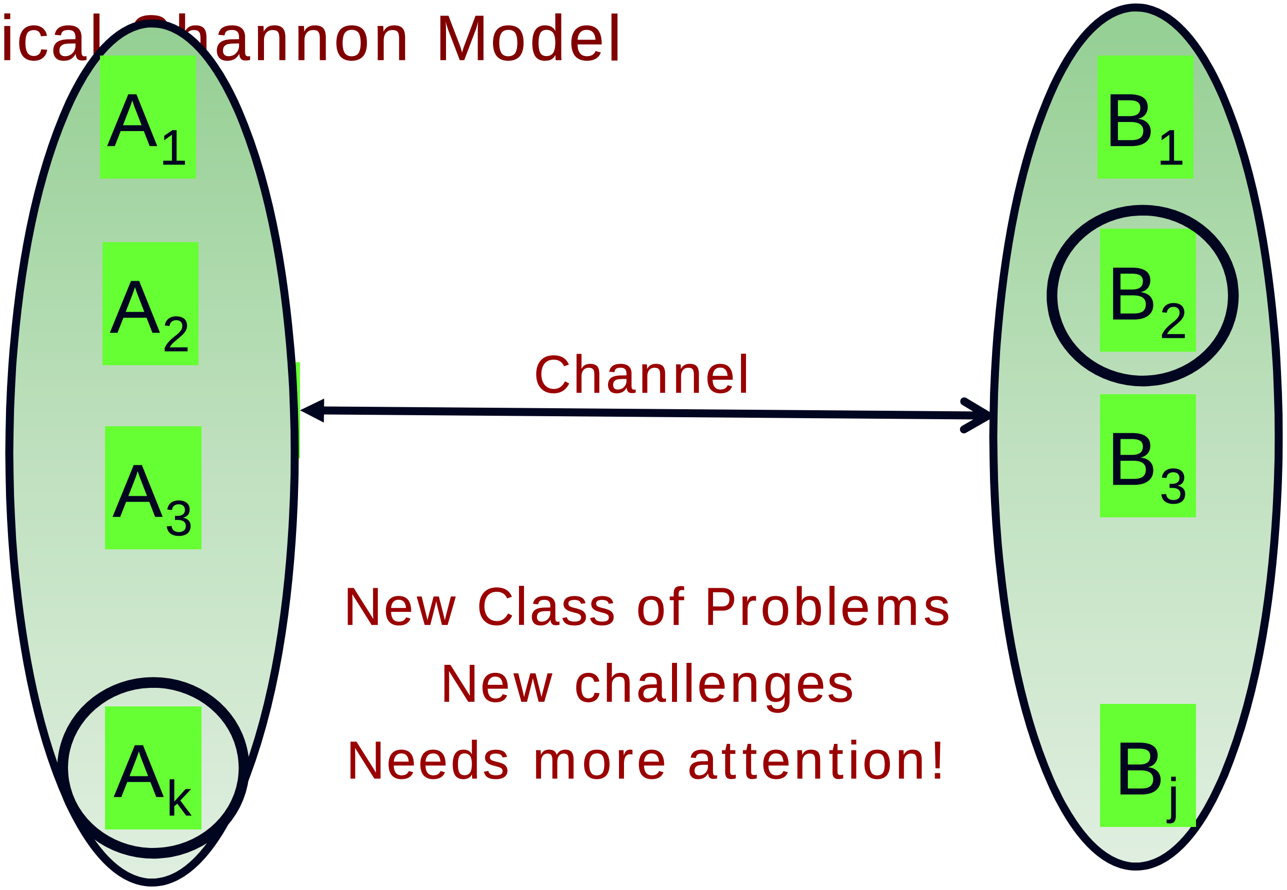
- Problem: Uncertainty of the future.
  - What formats/systems will prevail?
  - Why aren't software systems ever constant?

## Challenge:

- If Decoder does not know the Encoder, how should it try to guess what it meant?
- Similar example:
  - Learning to speak a foreign language
    - Humans do … (?)
      - Can we understand how/why?
      - Will we be restricted to talking to humans only?
      - Can we learn to talk to “aliens”? Whales? ☺
- Claim:
  - Questions can be formulated mathematically.
  - Solutions still being explored.

# Modelling uncertainty

Uncertain Communication Model
Classical Shannon Model



# Modern questions/answers

- Communicating players share large context.
  - Knowledge of English, grammar, socio-political context
  - Or ... Operating system, communication protocols, apps, compression schemes.
- But sharing is not perfect.
  - Can we retain some of the benefit of the large shared context, when sharing is imperfect?
  - Answer: Yes ... in many cases ... [ongoing work]
    - New understanding of human mechanisms
    - New reliability mechanisms coping with uncertainty!

## Language as compression

- Why are dictionaries so redundant+ambiguous?
  - Dictionary = map from words to meaning
  - For many words, multiple meanings
  - For every meaning, multiple words/phrases
  - Why?
- Explanation: “Context”
  - Dictionary:
    - Encoder: Context1 × Meaning → Word
    - Decoder: Context2 × Word → Meaning
    - Tries to compress length of word
    - Should works even if Context1 ≠ Context2
- [Juba,Kalai,Khanna,S’11],[Haramaty,S’13]: Can design encoders/decoders that work with uncertain context.

## Summary

- Reliability in Communication
  - Key Engineering problem of the past century
    - Led to novel mathematics
    - Remarkable solutions
    - Hugely successful in theory and practice
  - New Era has New Challenges
    - Hopefully new solutions, incorporating ideas from ...
      - Information theory, computability/complexity, game theory, learning, evolution, linguistics ...
    - ... Further enriching mathematics

# Thank You!

# A challenging special case

- Say Alice and Bob have rankings of N movies.
  - Rankings = bijections $\pi, \sigma : [N] \rightarrow [N]$
  - $\pi(i)$ = rank of $i^{\text{th}}$ player in Alice's ranking.
- Further suppose they know rankings are close.
  - $\forall\, i \in [N]: |\pi(i) - \sigma(i)| \leq 2$.
- Bob wants to know: Is $\pi^{-1}(1) = \sigma^{-1}(1)$
- How many bits does Alice need to send (non-interactively).
  - With shared randomness – $O(1)$
  - Deterministically?
    - $O(1)$? $O(\log N)$? $O(\log\log\log N)$?

## Meaning of Meaning?

- Difference between meaning and words
  - Exemplified in
    - Turing machine vs. universal encoding
    - Algorithm vs. computer program
  - Can we learn to communicate former?
    - Many universal TMs, programming languages
- [Juba,S.'08], [Goldreich,Juba,S.'12]:
  - Not generically …
  - Must have a goal: what will we get from the bits?
  - Must be able to sense progress towards goal.
  - Can use sensing to detect errors in understanding, and to learn correct meaning.
- [Leshno,S'13]:
  - Game theoretic interpretation

# Communication as Coordination Game [Leshno,S.'13]

- Two players playing series of coordination games
  - Coordination?
    - Two players simultaneously choose 0/1 actions.
    - "Win" if both agree:
      - Alice's payoff: not less if they agree
      - Bob's payoff: strictly higher if they agree.
    - How should Bob play?
      - Doesn't know what Alice will do. But can hope to learn.
      - Can he hope to eventually learn her behavior and (after finite # of miscoordinations) always coordinate?
- Theorem:
  - Not Deterministically (under mild "general" assumptions)
  - Yes with randomness (under mild restrictions)