Communication Amid Uncertainty

Madhu Sudan
Microsoft Research

Based on
- Goldreich, Juba, S. (JACM 2011)
- Juba, Kalai, Khanna, S. (ITCS 2011)
- Haramaty, S. (ITCS 2014)
- Canonne, Guruswami, Meka, S. (ITCS 2015)
- Leshno, S. (manuscript)
Communication vs. Computation

- Interdependent technologies: Neither can exist without other
- Technologies/Products/Commerce developed (mostly) independently.
  - Early products based on clean abstractions of the other.
  - Later versions added other capability as afterthought.
  - Today products ... deeply integrated.
- Deep theories:

Well separated ... and have stayed that way

Turing '36

Shannon '48
Consequences of the wall

- **Computing theory:**
  - Fundamental principle = Universality
  - You can program your computer to do whatever you want.
  - ⇒ Heterogeneity of devices

- **Communication theory:**
  - Centralized design (Encoder, Decoder, Compression, IPv4, TCP/IP).
  - You can NOT program your device!
  - ⇐ Homogeneity of devices

- **Contradiction! But does it matter?**
  - Yes!
Sample problems:

- Universal printing:
  - You are visiting a friend. You can use their Wifi network, but not their printer. Why?

- Projecting from your laptop:
  - Machines that learn to communicate, and learn to understand each other.

- Digital libraries:
  - Data that lives forever (communication across time), while devices change.
Essence of “semantics”: Uncertainty

- Shannon:
  - “The significant aspect is that the actual message is one selected from a set of possible messages”

- Essence of unreliability today:
  - Context: Determines set of possible messages.
    - dictionary, grammar, general knowledge
    - coding scheme, prior distribution, communication protocols ...
  - Context is HUGE; and not shared perfectly;
Modelling uncertainty

Uncertain Communication Model
Classical Shannon Model

New Class of Problems
New challenges
Needs more attention!
Hope

- Better understanding of existing mechanisms
  - In natural communication
  - In “ad-hoc” designs

- What problems are they solving?

- Better solutions?
  - Or at least understand how to measure the quality of a solution.
II: Uncertain Compression
Human-Human Communication

- Role of dictionary = ?
  - [Juba, Kalai, Khanna, S. 11]
- Dictionary: list of words representing message
  - words appear against multiple messages
  - multiple words per message.
- How to decide which word to use? Context!
  - Encoding: Given message, use shortest unambiguous word in current context.
  - Decoding: Given word, use most likely message in current context, (among plausible messages)
- Context = ???. Prob. distribution on messages
  \[ P_i = \text{Prob} [\text{message} = M_i] \]
Human Communication - 2

- Good (Ideal?) dictionary
  - Should compress messages to entropy of context:
    \[ H(P = \langle P_1, \ldots, P_N \rangle). \]

- Even better dictionary?
  - Should not assume context of sender/receiver identical!
  - Compression should work even if sender **uncertain** about receiver (or receivers’ context).

**Theorem [JKKS]:** If dictionary is “random” then compression achieves message length \( H(P) + \Delta \), if sender and receiver distributions are “\( \Delta \)-close”.

\[ M_1 = w_{11}, w_{12}, \ldots \]
\[ M_2 = w_{21}, w_{22}, \ldots \]
\[ M_3 = w_{31}, w_{32}, \ldots \]
\[ M_4 = w_{41}, w_{42}, \ldots \]

[Diagram of sender and receiver contexts with arrows indicating uncertainty.]
Implications

- Reflects tension between ambiguity resolution and compression.
  - Larger the gap in context ($\Delta$), larger the encoding length.
- Coding scheme reflects human communication?
- “Shared randomness” debatable assumption:
  - Dictionaries do have more structure.
  - Deterministic communication? [Haramaty+S,14]
  - Randomness imperfectly shared? Next ...
III: Imperfectly Shared Randomness
Communication (Complexity)

- Compression (Shannon, Noiseless Channel)
  
  $x \sim P = (P_1, \ldots, P_n)$

- What will Bob do with $x$?
  - Often knowledge of $x$ is overkill.
  - [Yao]’s model:
    - Bob has private information $y$.
    - Wants to know $f(x, y) \in \{0, 1\}$.
    - Can we get away with much less communication?

In general, model allows interaction. For this talk, only one way comm.
Brief history

- ∃ problems where Alice can get away with much fewer bits of communication.
  - Example: $\oplus (x, y) \triangleq \bigoplus_i (x_i \oplus y_i)$
  - But very few such deterministically.

- Enter Randomness:
  - Alice & Bob share random string $r$ (ind. of $x, y$)
  - Many more problems; Example:
    - $\text{Eq}(x, y) = 1$ if $x = y$ and 0 otherwise
      - Deterministically: $\Theta(n)$
      - Randomized: $O(1)$

- Uncertainty-motivated question:
  - Does randomness have to be perfectly shared?
Results

- [Newman ‘90s]:
  \[ CC \text{ without sharing} \leq CC \text{ with sharing} + \log n \]
- But additive cost of \( \log n \) may be too much.
  - Compression! Equality!!
- Model recently studied by [Bavarian et al.’14]
  - Equality: \( O(1) \) bit protocol w. imperfect sharing
- Our Results: [Canonne, Guruswami, Meka, S.’15]
  - Compression: \( O(H(P) + \Delta) \)
  - Generally: \( k \) bits with shared randomness
    \[ \Rightarrow 2^k \text{ bits with imperfect sharing.} \]
  - \( k \to 2^k \) loss is necessary.
Some General Lessons

- Compression Protocol:
  - Adds “error-correction” to [JKKS] protocol.
  - Send shortest word that is far from words of other high probability messages.
  - Another natural protocol.

- General Protocol:
  - Much more “statistical”
    - Classical protocol for Equality:
      - Alice sends random coordinate of ECC(x)
    - New Protocol
      - ~ Alice send # 1’s in random subset of coordinates.
IV: Coordination
Communicate meaning?

- Ultimate goal:
  - Message $\Rightarrow$ Instructions.
    - What is this dictionary?
    - Can it be learned by communication?
- At first glance:
  - Ambiguity can never be resolved by communication (even a theorem [JS’08]).
- Second look:
  - Needs more careful definitions.
    - Meaning = mix of communication + actions + incentives.
(Mis) Understanding?

- Uncertainty problem:
  - Sender/receiver disagree on meaning of bits

- Definition of Understanding?
  - Sender sends instructions; Receiver follows?
    - Errors undetectable (by receiver)
    - Not the right definition anyway:
      - Does receiver want to follow instructions
      - What does receiver gain by following instructions? Must have its own “Goal”/”Incentives”.

- [Goldreich, Juba, S. 2012]: Goal-oriented communication:
(Mis) Understanding?

- Uncertainty problem:
  - Sender/receiver disagree on meaning of bits

- Definition of Understanding?
  - Receiver has goals/incentives.

- [Goldreich, Juba, S. 2012]: Goal-oriented communication:
  - Define general communication problems (and goals)
  - Show that if
    - Sender can help receiver achieve goal (from any state)
    - Receiver can sense progress towards goal
  - then
    - Receiver can achieve goal.

- Functional definition of understanding.
Illustration: (Repeated) Coordination

- [Leshno, S.]
- Basic Coordination Game:
  - Alice and Bob simultaneously choose actions $\in \{0,1\}$
  - If both pick same action, both win.
  - If they pick opposite actions, both lose.
- Main challenge: Don’t know what the other person will choose when making our choice.
- Repeated version:
  - Play a sequence of games, using outcome of previous games to learn what the other player may do next.
  - Goal: Eventual perpetual coordination.
Our setting

- Repeated coordination game with uncertainty:
  - Bob’s perspective:
    - Knows his payoffs – 1 for coord.; 0 for not.
    - Does not know Alice’s payoffs (uncertainty):
      - May vary with round
      - But for every action of Alice, payoff does not decrease if Bob coordinates (compatibility).
      - Knows a set $S_A$ of strategies she may employ ("reasonable behaviors").
  - Can he learn to coordinate eventually?
Coordination with Uncertainty

- Mixes essential ingredients:
  - Communication: Actions can be used to communicate (future actions).
  - Control: Communication (may) influence future actions.
  - Incentives:
    - Bob has incentive to coordinate.
    - Alice not averse.

- What do the general results say?
  - \( \exists \) Universal strategy \( U \) s.t.
    - \( \forall \) Alice s.t. \( \exists \) Bob who coordinates with Alice from any state.
    - \( U \) coordinates with Alice.
Lessons

- Coordination is possible:
  - Even in extreme settings where
    - Alice has almost no idea of Bob
    - Bob has almost no idea of Alice
    - Alice is trying to learn Bob
    - Bob is trying to learn Alice
  - Learning is slow ...
    - Need to incorporate beliefs to measure efficiency. [Juba, S. 2011]
    - Does process become more efficient when languages have structure? [Open]
Conclusions

- Context in communication:
  - White elephant in the room.
    - Huge, unmentionable, weighing us down.
  - Context usually imperfectly shared.
  - Uncertainty + Scale $\Rightarrow$ New class of problems.
  - What are new “error-correcting” mechanisms?
    - Can be build reliability on top of unreliability?
Thank You!