Communication amid Uncertainty

Madhu Sudan
Microsoft, Cambridge, USA

Based on:

- Goal-Oriented Communication – Goldreich, Juba & S. (JACM 2012)
- Compression without a common prior ... – Kalai, Khanna, Juba & S. (ICS 2011)
- Efficient Semantic Communication with Compatible Beliefs – Juba & S. (ICS 2011)
- Deterministic Compression with uncertain priors – Haramaty & S. (ITCS 2014)
Classical theory of communication

Shannon (1948)

- Clean architecture for reliable communication.

- Remarkable mathematical discoveries: Prob. Method, Entropy, (Mutual) Information

- Needs reliable encoder + decoder (two reliable computers).
Uncertainty in Communication?

- Always has been a central problem:
  - But usually focusses on uncertainty introduced by the channel
- Standard Solution:
  - Use error-correcting codes
  - Significantly:
    - Design Encoder/Decoder jointly
    - Deploy Encoder at Sender, Decoder at Receiver
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 1

- Printing in a new environment
  - Say, you are visiting a new university.
  - Printer is intelligent; so is your computer;
    - Can’t they figure out how to talk to each other?
- Problem (with current designs):
  - Computers need to know about the printer already to print on them.
  - Why can’t they also figure out how future printers will work?
    - Uncertainty (about printers of the future).
Example 2

- 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 systems will prevail?
  - Why aren’t software systems ever constant?
Modelling uncertainty

Semantic Communication Model

Classical Shannon Model

New Class of Problems
New challenges
Needs more attention!
Nature of uncertainty

- $A_i$’s, $B_j$’s differ in beliefs, but can be centrally programmed/designed.
  - [Juba, Kalai, Khanna, S.’11]: Compression in this context has graceful degradation as beliefs diverge.
  - [Haramaty, S’13]: Role of randomness in this context.

- $A_i$’s, $B_j$’s differ in behavior:
  - Nothing to design any more (behavior already fixed).
  - Best hope: Can identify certain $A_i$’s (universalists) that can interact successfully with many $B_j$’s. Can eliminate certain $B_j$’s on the grounds of “limited tolerance”.
  - [Juba, S’08; Goldreich, J, S’12; J, S’11]: “All is not lost, if we keep goal of communication in mind”
  - [Leshno, S’13]: “Communication is a Coordination Game”
  - Details don’t fit in margin ...
II: Compression under uncertain beliefs/priors
Motivation

- New era of challenges needs new solutions.
  - Most old solutions do not cope well with uncertainty.
  - The one exception?
    - Natural communication (Humans ↔ Humans)
- What are the rules for human communication?
  - “Grammar/Language”
  - What kind of needs are they serving?
  - What kind of results are they getting? (out of scope)
  - If we were to design systems serving such needs, what performance could they achieve?
Role of Dictionary (/Grammar/Language)

- Dictionary: maps words to meaning
  - Multiple words with same meaning
  - Multiple meanings to same word
- How to decide what word to use (encoding)?
- How to decide what a word means (decoding)?
  - Common answer: Context
- Really Dictionary specifies:
  - Encoding: context $\times$ meaning $\rightarrow$ word
  - Decoding: context $\times$ word $\rightarrow$ meaning
- Context implicit; encoding/decoding works even if context used not identical!
Context?

- In general complex notion ...
  - What does sender know/believe
  - What does receiver know/believe
  - Modifies as conversation progresses.

- Our abstraction:
  - Context = Probability distribution on potential “meanings”.
  - Certainly part of what the context provides; and sufficient abstraction to highlight the problem.
The problem

- Wish to design encoding/decoding schemes (E/D) to be used as follows:
  - Sender has distribution $P$ on $M = \{1, 2, \ldots, N\}$
  - Receiver has distribution $Q$ on $M = \{1, 2, \ldots, N\}$
  - Sender gets $X \in M$
  - Sends $E(P, X)$ to receiver.
  - Receiver receives $Y = E(P, X)$
  - Decodes to $\hat{X} = D(Q, Y)$

- Want: $X = \hat{X}$ (provided $P, Q$ close),
  - While minimizing $\text{Exp}_{X \leftarrow P} |E(P, X)|$
Closeness of distributions:

- \( P \) is \( \Delta \)-close to \( Q \) if for all \( X \in M \),
  \[
  \frac{1}{2^\Delta} \leq \frac{P(X)}{Q(X)} \leq 2^\Delta
  \]

- \( P \) \( \Delta \)-close to \( Q \) \( \Rightarrow \) \( D(P||Q), D(Q||P) \leq \Delta \).
Dictionary = Shared Randomness?

- Modelling the dictionary: What should it be?

- Simplifying assumption – it is shared randomness, so ...

- Assume sender and receiver have some shared randomness $R$ and $X, P, Q$ independent of $R$.
  - $Y = E(P, X, R)$
  - $\hat{X} = D(Q, Y, R)$

- Want $\forall X, \Pr_{R}[\hat{X} = X] \geq 1 - \epsilon$
Solution (variant of Arith. Coding)

- Use R to define sequences
  - $R_1 [1], R_1 [2], R_1 [3], ...$
  - $R_2 [1], R_2 [2], R_2 [3], ...$
  - ... 
  - $R_N [1], R_N [2], R_N [3], ....$

- $E_\Delta(P, x, R) = R_x[1 ... L], \text{ where } L \text{ chosen s.t. } \forall z \neq x$
  - Either $R_z[1 ... L] \neq R_x[1 ... L]$
  - Or $P(z) < \frac{P(x)}{4^\Delta}$

- $D_\Delta(Q, y, R) = \arg\max_{\hat{x}} \{Q(\hat{x})\} \text{ among } \hat{x} \in \{z \mid R_z[1 ... L] = y\}$
Performance

- Obviously decoding always correct.

- Easy exercise:
  - \( \text{Exp}_X [E(P,X)] = H(P) + 2 \Delta \)

- Limits:
  - No scheme can achieve \( (1 - \epsilon) \cdot [H(P) + \Delta] \)
  - Can reduce randomness needed.
Implications

- Reflects the tension between ambiguity resolution and compression.
  - Larger the \( \Delta \) ((estimated) gap in context), larger the encoding length.
  - Entropy is still a valid measure!
- Coding scheme reflects the nature of human process (extend messages till they feel unambiguous).
- The “shared randomness” assumption
  - A convenient starting point for discussion
  - But is dictionary independent of context?
  - This is problematic.
III: Deterministic Communication Amid Uncertainty
A challenging special case

- Say Alice and Bob have rankings of \( N \) players.
  - 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?
    - \( O(1)? \ O(\log N)? \ O(\log \log \log \log N)? \)
Model as a graph coloring problem

- Consider family of graphs $U_{N,\ell}$:
  - Vertices = permutations on $[N]$
  - Edges = $\ell$-close permutations with distinct messages. (two potential Alices).

- Central question: What is $\chi(U_{N,\ell})$?
Main Results [w. Elad Haramaty]

- Claim: Compression length for toy problem
  \[ \in \left[ \log \chi(U_{N,2}), \log \chi(U_{N,4}) \right] \]

- Thm 1: \[ \chi(U_{N,\ell}) \leq \ell^O(\ell \log^* N) \]
  - \[ \log^{(i)} N \equiv \log \log \ldots N \ (i \ \text{times}) \]
  - \[ \log^* N \equiv \min \{ i \mid \log^{(i)} N \leq 1 \} \].

- Thm 2: \exists\ \text{uncertain comm. schemes with}
  1. \[ \text{Exp}_m[|E(P, m)|] \leq O(H(P) + \Delta + \log \log N) \]
     \hspace{1cm} (0\text{-error}).
  2. \[ \text{Exp}_m[|E(P, m)|] \leq \ell^O(\varepsilon^{-1}(H(P) + \Delta + \log^* N)) \]
     \hspace{1cm} (\varepsilon \text{-error}).

- Rest of the talk: Graph coloring

12/02/2013 Purdue: Uncertainty in Communication
Restricted Uncertainty Graphs

- Will look at $U_{N,\ell,k}$
  - Vertices: restrictions of permutations to first $k$ coordinates.
  - Edges: $\pi' \leftrightarrow \sigma'$
    $\iff \exists \pi$ extending $\pi'$ and $\sigma$ extending $\sigma'$ with $\pi \leftrightarrow \sigma$
Homomorphisms

- $G$ homomorphic to $H$ ($G \to H$) if
  \[ \exists \phi : V(G) \to V(H) \text{ s.t. } u \leftrightarrow_G v \Rightarrow \phi(u) \leftrightarrow_H \phi(v) \]

- Homomorphisms?
  - $G$ is $k$-colorable $\iff G \to K_k$
  - $G \to H$ and $H \to L \Rightarrow G \to L$

- Homomorphisms and Uncertainty graphs.
  - $U_{N,\ell} = U_{N,\ell,N} \to U_{N,\ell,N-1} \to \ldots \to U_{N,\ell,\ell+1}$
  - Suffices to upper bound $\chi(U_{N,\ell,k})$
Chromatic number of $U_{N,\ell,\ell+1}$

- For $f: [N] \rightarrow [2\ell]$, Let
  \[ I_f = \{ \pi \mid f(\pi_1) = 1, f(\pi_i) \neq 1, \forall i \in [2\ell] - \{1\} \} \]

- Claim: $\forall f, I_f$ is an independent set of $U_{N,\ell,\ell+1}$

- Claim: $\forall \pi, \Pr_f[\pi \in I_f] \geq \frac{1}{4\ell}$

- Corollary: $\chi(U_{N,\ell,\ell+1}) \leq O(\ell^2 \log N)$
Better upper bounds:

- Say $\phi: G \to H$

- $d_\phi(u) \equiv |\{ \phi(v) | v \leftrightarrow_G u \}|$
  
  $d_\phi \equiv \max_u \{d_\phi(u)\}$

- Lemma:
  
  $\chi(G) \leq O(d_\phi^2 \log \chi(H))$

- For $\phi_k: U_{N,\ell,k} \to U_{N,\ell,k-\ell}$
  
  $d_{\phi_k} = \ell^{O(k)}$
Better upper bounds:

- \( d_\phi \equiv \max_u \{|\phi(v)|v \leftrightarrow_G u|\} \)

- **Lemma:** \( \chi(G) \leq O(d_\phi^2 \log \chi(H)) \)

- For \( \phi_k: U_{N,\ell,k} \rightarrow U_{N,\ell,k-\ell}, \quad d_{\phi_k} \leq \ell^{O(k)} \)

- **Corollary:** \( \chi(U_{N,\ell,k}) \leq \ell^{O(k)} \log^{\frac{k}{\ell}} N \)

- **Aside:** Can show: \( \chi(U_{N,\ell,k}) \geq \log^{\Omega(\frac{k}{\ell})} N \)
  - Implies can’t expect simple derandomization of the randomized compression scheme.
Future work?

- **Open Questions:**
  - Is $\chi(U_{N,\ell}) = O_\ell(1)$?
  - Can we compress arbitrary distributions to $O(H(P) + \Delta)$? $O(H(P) + \Delta + \log^* N)$? or even $O(H(P) + \Delta + \log \log \log N)$?

- **On conceptual side:**
  - Better understanding of forces on language.
    - Information-theoretic
    - Computational
    - Evolutionary
    - Game-theoretic
  - Design better communication solutions!
Thank You