Context in Communication

Madhu Sudan
Harvard University

Based on conversations with Brendan Juba, Oded Goldreich, Elad Haramaty, Badih Ghazi, Boriana Gjura

Nov. 9, 2019 Shafifest@FOCS: Context in Communication 1 of 20
This Talk

- Context & Motivation
- Some past works
- Some challenges for future
Context (in Communication) = ?

- Empirically, Informally:
  - Huge piece of information (much larger than “content” of communication)
  - Not strictly needed for communication ...
  - ... But makes communication efficient, when shared by communicating players
  - ... helps even if context not shared perfectly.
- Examples: Language, General knowledge, Math
- Challenge: Formalize? ...
- ... but first some motivation
Sales Pitch + Intro

- Most of communication theory [a la Shannon, Hamming]:
  - Built around sender and receiver perfectly synchronized.
  - So large context (codes, protocols, priors) ignored.
- Most Human communication (also device-device)
  - ... does not assume perfect synchronization.
  - So context is relevant:
    - Qualitatively (receiver takes wrong action)
    - and Quantitatively
Ingredients in Human Communication

- Ability to start with (nearly) zero context and "learning to communicate by communicating".
  - Children learn to speak ... (not by taking courses)
  - Focus of work with Brendan and Oded
  - "What is possible?" (a la computability)

- Ability to leverage large uncertain context.
  - E.g., ... This talk today ...
  - Assumes ... English, Math, TCS, Social info, Geography.

- Aside ... what is "self-contained"?

- "How to make communication efficient (using context)?" (a la complexity)
The model (with shared randomness)

The function $f: (x, y) \mapsto \Sigma$ is usually studied for lower bounds. This talk: CC as +ve model.

$R = $$$

$CC(f) = \# \text{bits exchanged by best protocol}$

$f(x, y) \text{ w.p. } 2/3$
Communication & (Uncertain) Context

Context_{Alice} \equiv (f_A, R_A, P_{XY}^A)

Context_{Bob} \equiv (f_B, R_B, P_{XY}^B)

f : (x, y) \mapsto \Sigma

x \rightarrow Alice \rightarrow y

Bob \rightarrow f(x, y) \text{ w.p. } 2/3
Part II: Some recent works
1. Imperfectly Shared Randomness

Context_{Alice} \equiv R_A

\text{Context}_{Bob} \equiv R_B

f: (x, y) \mapsto \Sigma

\begin{align*}
\text{Alice} & \quad x \\
& \quad f(x, y) \quad \text{w.p. } \frac{2}{3}
\end{align*}

\begin{align*}
\text{Bob} & \quad y \\
& \quad f(x, y)
\end{align*}

\text{Results:}

[Bavarian, Gavinsky, Ito]:
- Equality has $O(1)$ CC.

[Cannone, Guruswami, Meka, S]:
- Imperfection blows up communication exponentially

- $k$ bits $\to O(2^k)$ bits;
- This is necessary (for large $n$)
2. Uncertain Functionality

Context_{Alice} \equiv (f, R, P_{X|Y})

Context_{Bob} \equiv (f, R, P_{X|Y})

\[ f : (x, y) \mapsto \sum \text{ w.p. } \frac{2}{3} \]

\[ f \sim P_{X|Y} \]

[Ghazi, Komargodksi, Kothari, S]:
- Gave precise model
- 1-bit cc problem where uncertainty leads to \( \Omega(\sqrt{n}) \) communication
- Doesn’t happen if \( x \perp y \)

Moral: Protocol is not a continuous function of \( f \)
3. Compression & (Uncertain) Context

Context \text{Alice} = P^A_x \\

Context \text{Bob} = P^B_x \\

\( x \rightarrow f: (x, y) \mapsto \Sigma \rightarrow y \)

\begin{align*}
[\text{Juba}, \text{Kalai}, \text{Khanna}, \text{S}], [\text{Haramaty}, \text{S}], \\
[\text{Ghazi}, \text{Haramaty}, \text{Kamath}, \text{S}]:
\end{align*}

- If \( P^A_x \approx P^B_x \) and \( \exists \) shared randomness then \( \mathbb{E}[\text{communication}] \approx H(P^A_x) \)
- Deterministically \( \leq H(P^A_x) + \log n \)

Relevance: Maybe a good analogy to languages

w.p. 2/3
Part III: Future?
Well understood ... (thanks to Shafi & co.)

- Interactive Proofs
- Zero Knowledge Proofs
- Multi-Prover Interactive Proofs
- PCPs
- Interactive Proofs for Muggles
- Pseudodeterministic Proofs ...

... nevertheless some challenges in understanding communication of proofs ...
Standard Assumption

- **Small (Constant) Number of Axioms**
  - \( X \rightarrow Y, Y \rightarrow Z \Rightarrow X \rightarrow Z, \text{ Peano, etc.} \)

- **Medium Sized Theorem:**
  - \( \forall x, y, z, n \in \mathbb{N}, \quad x^n + y^n = z^n \rightarrow n \leq 2 \ldots \)

- **Big Proof:**
  - Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah
The truth

- Mathematical proofs assume large context.
  - "By some estimates a proof that $2+2=4$ in ZFC would require about $20000$ steps ... so we will use a huge set of axioms to shorten our proofs – namely, everything from high-school mathematics"
    - [Lehman,Leighton,Meyer – Notes for MIT 6.042]
- Context (= huge set of axioms) shortens proofs.
- But context is uncertain!
  - What is “high school mathematics”? 

Nov. 9, 2019  Shafifest@FOCS: Context in Communication
Communicating ("speaking of") Proofs

Ingredients:

- **Prover:**
  - Axioms $A$, Theorem $T$, Proof $\Pi$
  - Communicates $(T, \Pi)$ (Claim: $\Pi$ proves $A \to T$)

- **Verifier:**
  - Complete+sound: $\exists \Pi, V^A(T, \Pi) = 1$ iff $A \to T$
  - Verifier efficient = $\text{Poly}(T, \Pi)$ with oracle access to $A$

- Axioms = Context.
Uncertainty of Context?

- **Prover:** works with axioms $A_p$, generates $\Pi$ s.t. $V^{A_p}(T, \Pi) = 1$
- **Verifier:** works with axioms $A_v$, checks $V^{A_v}(T, \Pi) = 1$
- **Robust prover/verifier?**
  - Need measure $\delta(A_p, A_v)$ (not symmetric).
  - Given $\delta$ prover should be able to generate $\Pi_\delta$ such that $\forall A_p \text{ s.t. } \delta(A_p, A_v) \leq \delta, V^{A_v}(T, \Pi) = 1$
  - $\Pi_\delta$ not much larger than $\Pi = \Pi_0$
  - $\delta(.,..)$ “reasonable”...
    - E.g. if $X \rightarrow Y, Y \rightarrow Z \in A$ and $A' = A \cup \{X \rightarrow Z\}$
      then $\delta(A', A)$ tiny?
- **Open:** Does such an efficient verifier exist?
Communication as a lens: Modelling the mind

- Mind does not just store and retrieve information.
- Also does reasoning.
- Challenge: Propose a model that explains its intricacies.
- Specifically, deals with "updates: Add "X" to knowledge updates: Add "X" to knowledge"
- "queries: Answers questions "Is Y true?" or "prove Y!"
- Some "intricacies"
- "Crossword Problem"

Proof:
We first transform 1 into a statement concerning boolean functions. Associate with the subgraph $G$ a boolean function $g$ such that $g(x) = 1$ iff $x \in V(G)$. Note that $\deg_G(x) = n - s(g, x)$ for $x \in V(G)$ and the same holds in $Q_n - G$ for $x \notin V(G)$. Denote by $E(g)$ the average value of $g$ on $C^n$. Now 1 and 2 are clearly equivalent to the following:

1. For any boolean function $g$, $E(g) \neq 0$ implies $\exists x: s(g, x) \leq n - h(n)$.
2. For any boolean function $f$, $s(f) < h(n)$ implies $d(f) < n$.

To see the equivalence of 1' and 2', define

$$g(x) = f(x) p(x),$$

where $p(x)$ is the parity function of $x$: $p(x) = \prod_{i=1}^{n} x_i$. Note that for all $x \in C^n$, $s(g, x) = n - s(f, x)$ and for all $I \subseteq [n]$, $\hat{g}(I) = \hat{f}([n] - I)$, therefore $E(g) = \hat{g}(\emptyset) = \hat{f}([n])$, where $\hat{f}([n])$ is the Fourier transform of $f$ at $[n]$, i.e., the highest order coefficient in the representation of $f$ as a polynomial.

1' $\Rightarrow$ 2'. Assume that $d(f) = n$, i.e., $\hat{f}([n]) \neq 0$. This is equivalent to $E(g) \neq 0$. By 1', $\exists x: s(g, x) \leq n - h(n)$; therefore $\exists x: s(f, x) \geq h(n)$, contradicting the premise.

2' $\Rightarrow$ 1'. Assume that $\forall x$, $s(g, x) > n - h(n)$. This implies that $s(f) < h(n)$. By 2', $d(f) < n$, which is equivalent to $\hat{f}([n]) = \hat{g}(\emptyset) = E(g) = 0$, contradicting the premise. $\blacksquare$

Crossword Problem
E.g.
Q. "Woodstock performer; J _ _ _ _ _ _ Z"
A. Don't know!
Q. "Woodstock performer; J O A _ _ _ _ Z"
A. Joan Baez!

Real question: Why couldn't I answer it first time?

"Crossword Problem"

Proof verification

Oct. 31, 2019  EPFL - Communication Amid Uncertainty
Conclusions

- Very poor understanding of “communication of proofs” as we practice it.
  - Have to rely on verifier’s “knowledge”
  - But can’t expect to know it exactly
- Exposes holes in computational understanding of knowledge-processing.
  - Can we “verify” more given more time?
  - Or are we just memorizing?
Thank You!

Happy Birthday, Shafi!!