Context in Communication

Madhu Sudan
Harvard University

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

This Talk

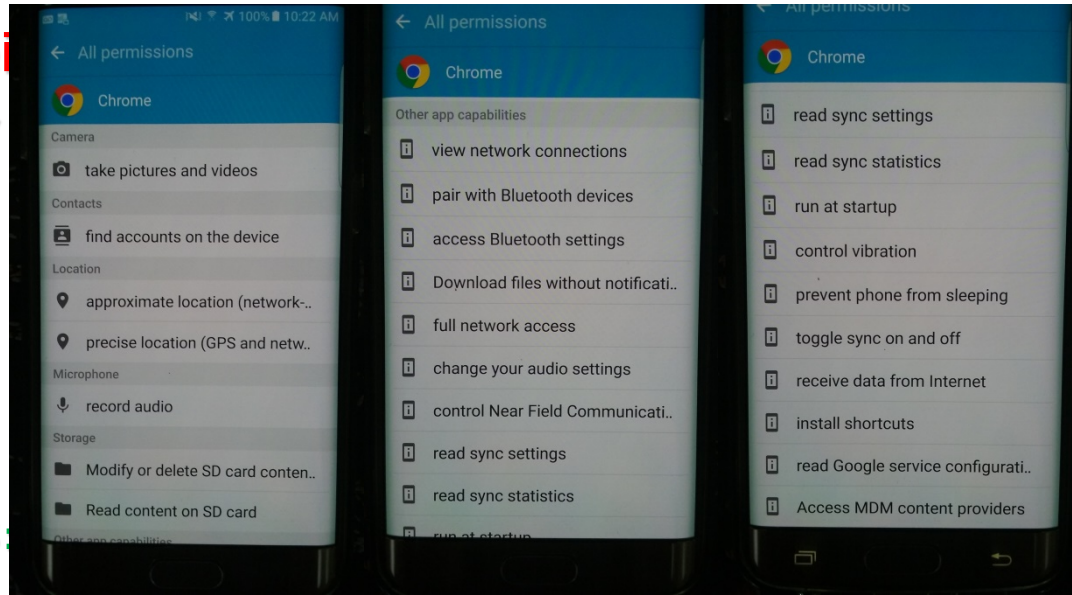
- Context & Motivation
- Some past works
- Some challenges for future

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

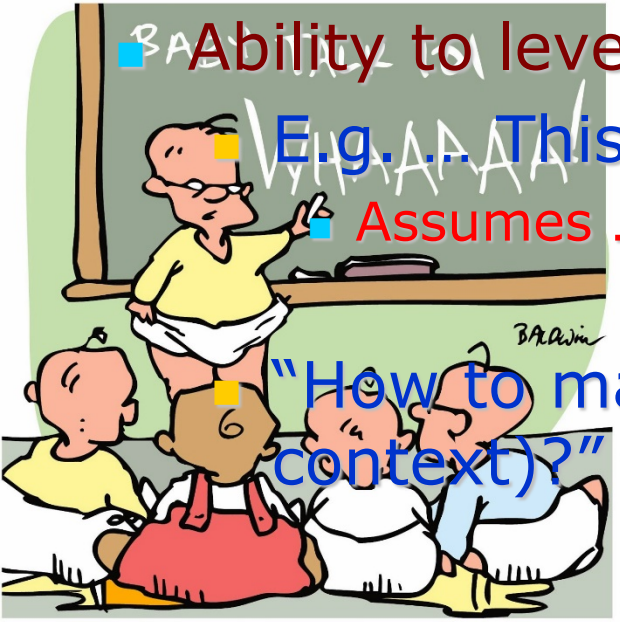
Sales Pitch + Intro

- 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 (recei
    - 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)

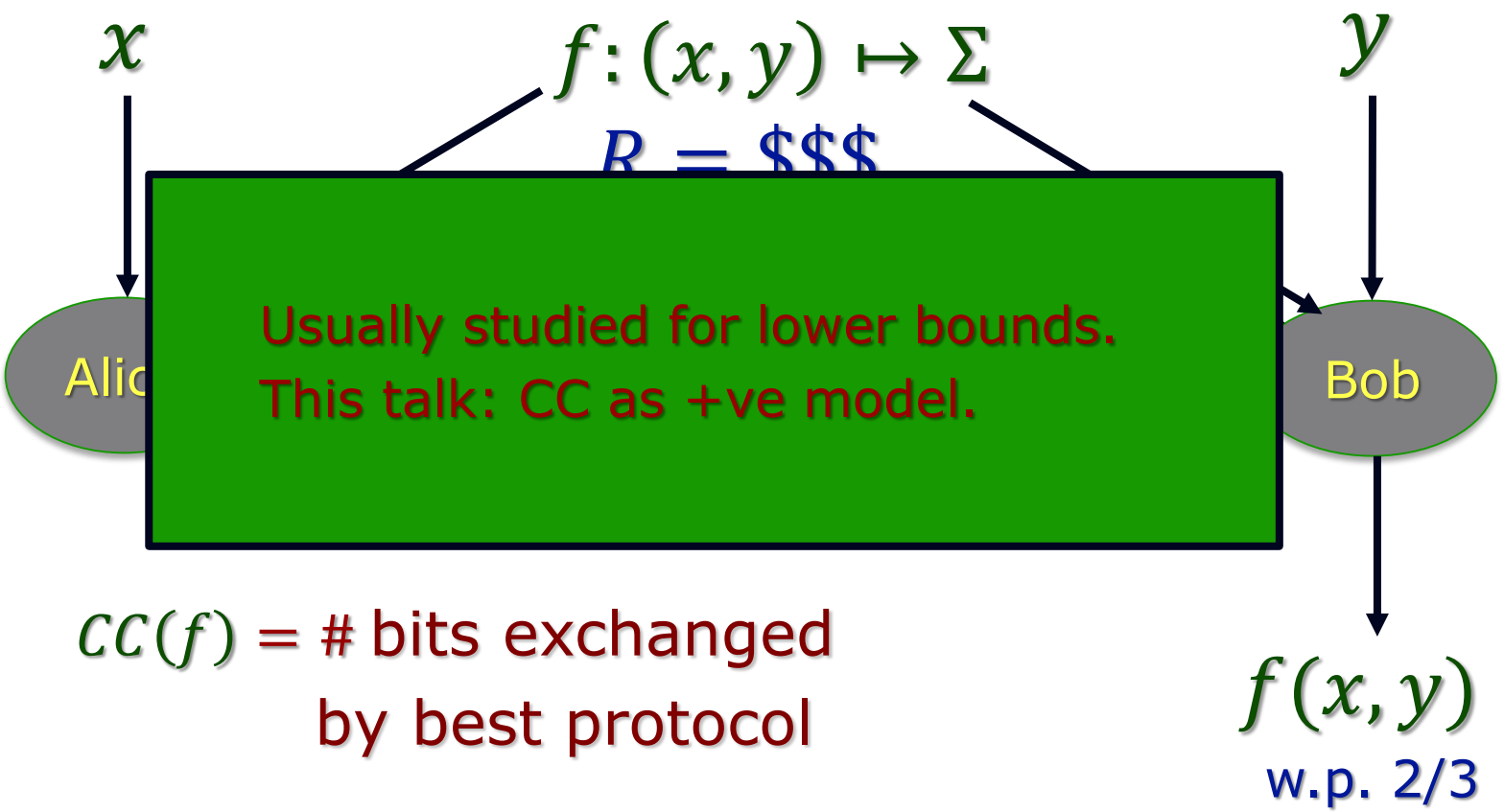


- Ability to leverage large uncertain context.
  - E.g. … This talk today …
    - Assumes … English, Math, TCS, Social info, Geography.
      - Aside … what is "self-contained"?
  - "How to make communication efficient (using context)?" (a la complexity)

"This is a good one. It means, 'Until my every need is met, your life will be hell.'"

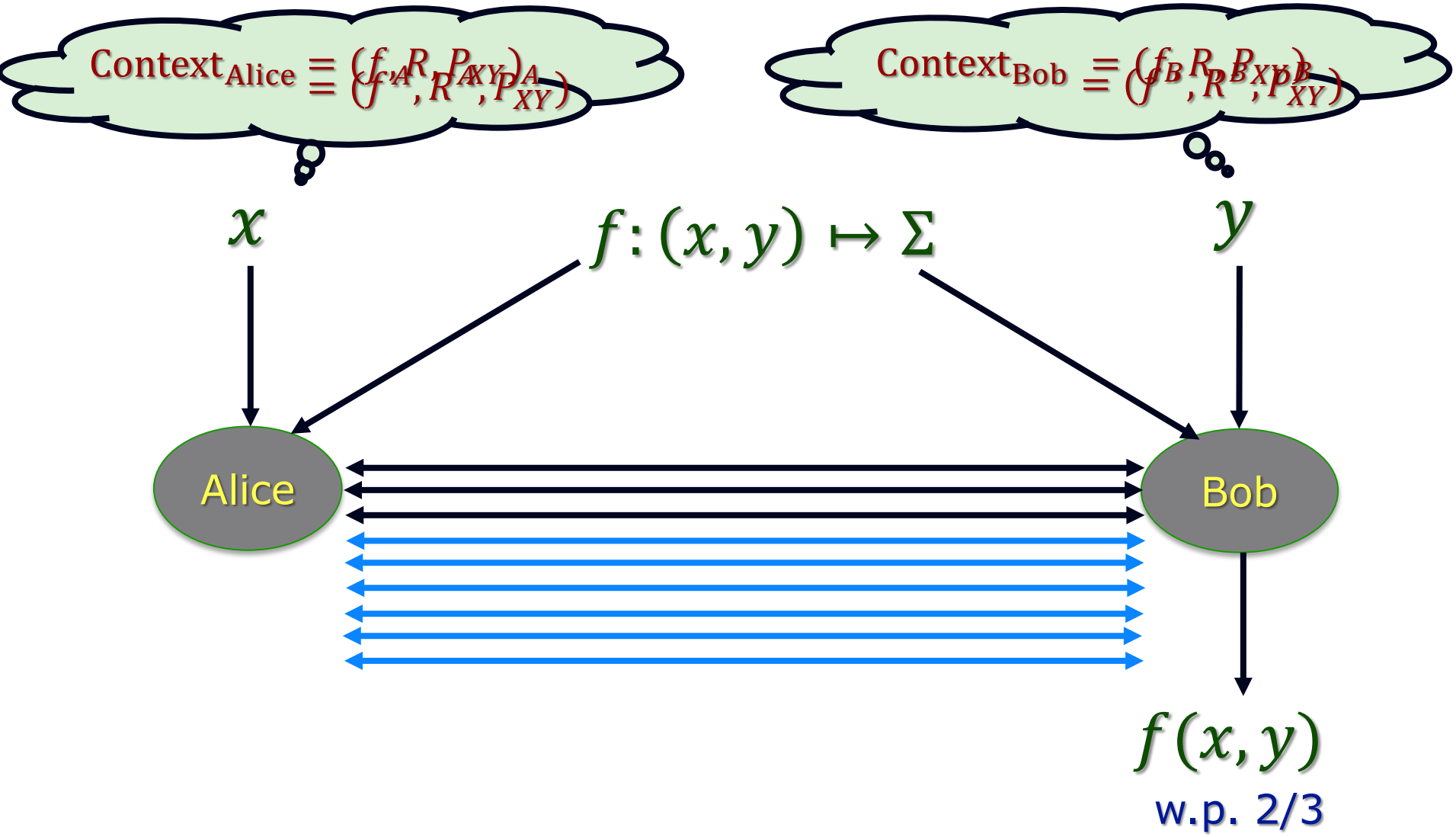
Underlying Model: Communication Complexity

The model (with shared randomness)



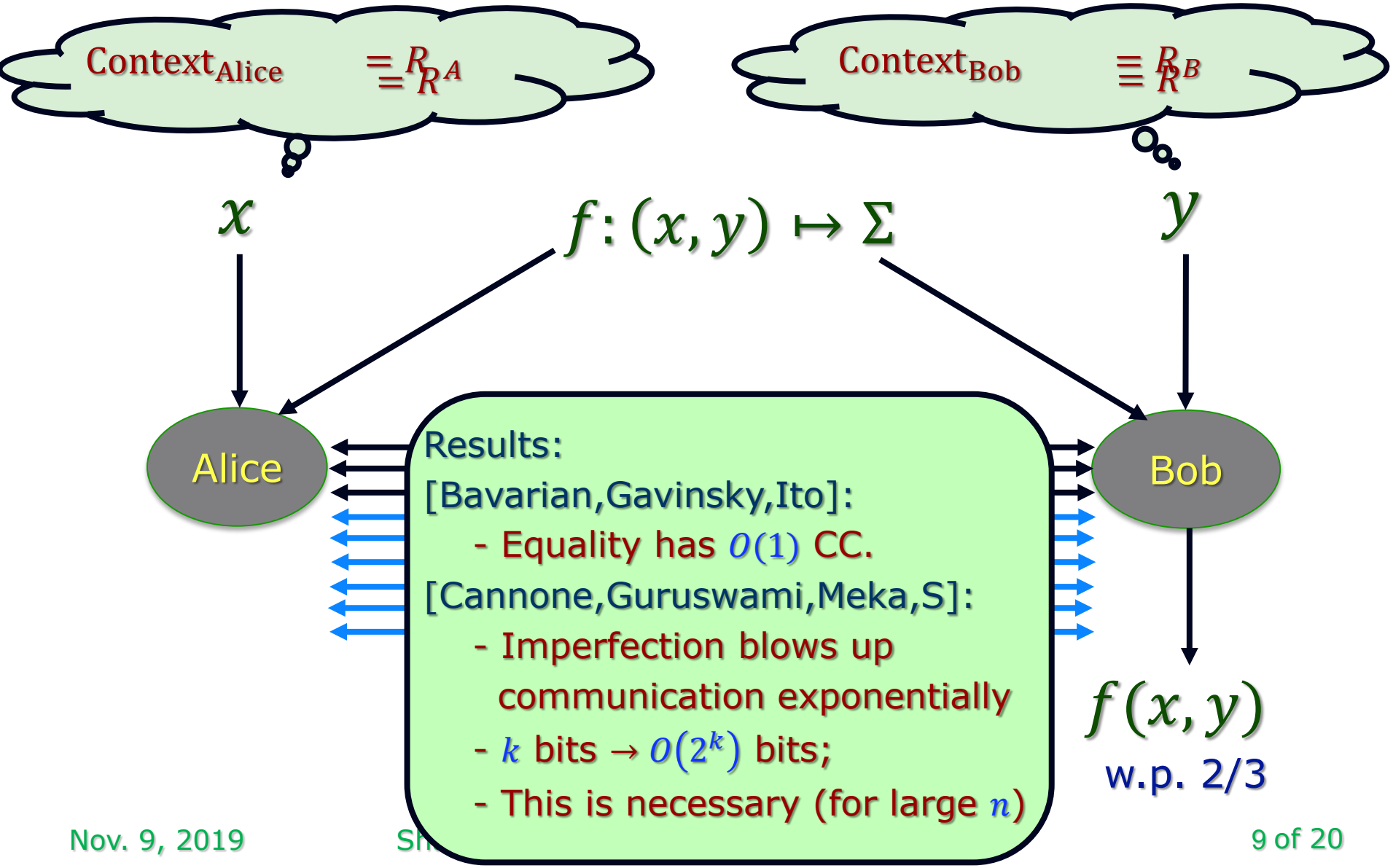
$CC(f)$ = # bits exchanged by best protocol

Communication & (Uncertain) Context

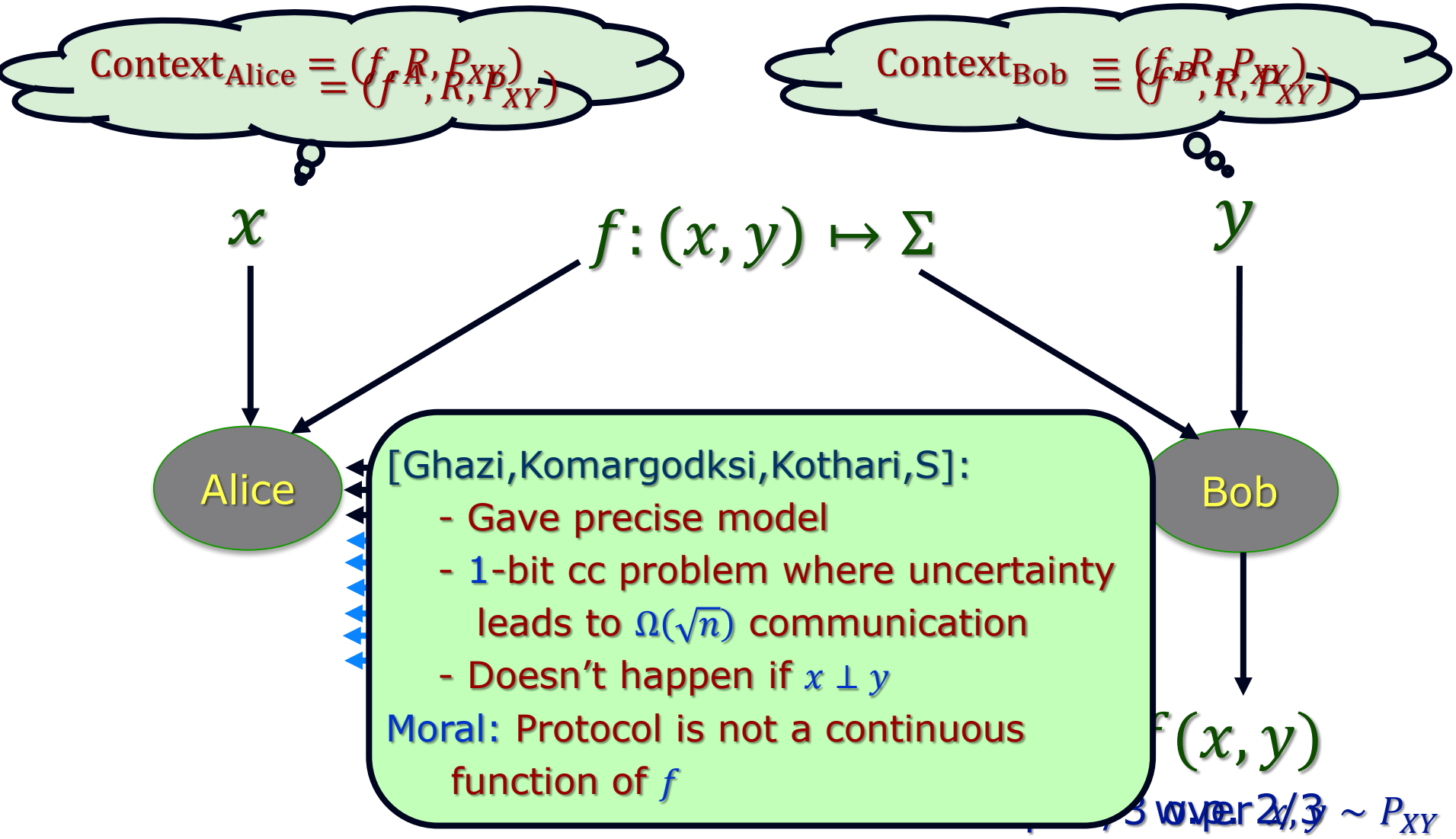


Part II: Some recent works

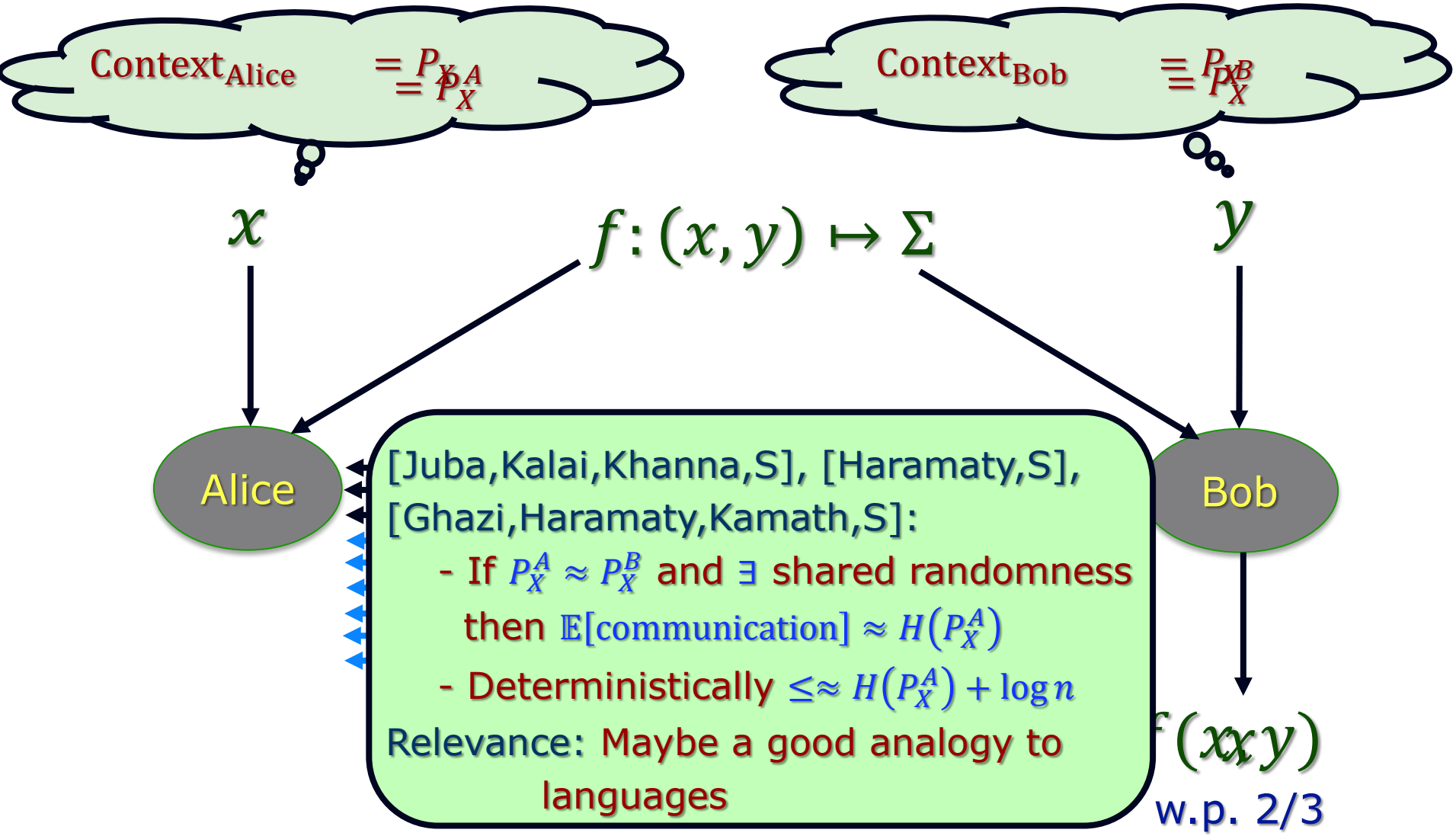
1. Imperfectly Shared Randomness



2. Uncertain Functionality



3. Compression & (Uncertain) Context



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

A	T	Π

- Small (Constant) Number of Axioms
  - $X \to Y, Y \to Z \Rightarrow \ X \to Z$, Peano, etc.
- Medium Sized Theorem:
  - $\forall x, y, z, n \in \mathbb{N}, \quad x^n + y^n = z^n \to n \leq 2$ …
- Big Proof:
  - Blah blah blah blah blah blah bla blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah 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"?

Communicating ("speaking of") Proofs

- Ingredients:

  - Prover:
    - Axioms A, Theorem T, Proof Π
    - Communicates (T, Π) (Claim: Π proves $A \to T$)
  - Verifier:
    - Complete+sound: $\exists \Pi, V^A(T, \Pi) = 1$ iff $A \to T$
    - Verifier efficient = Poly(T, Π) with oracle access to A
- Axioms = Context.

Uncertainty of Context?

- Prover: works with axioms A_P, generates Π 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 δ prover should be able to generate Π_δ such that $\forall\, A_P$ s.t. $\delta(A_P, A_V) \leq \delta,\ V^{A_V}(T, \Pi) = 1$
  - Π_δ not much larger than $\Pi = \Pi_0$
  - $\delta(.,.)$ "reasonable" …
    - E.g. if $X \to Y, Y \to Z \in A$ and $A' = A \cup \{X \to Z\}$ then $\delta(A', A)$ tiny?
- Open: Does such an efficient verifier exist?

Communication as a lens: Modelling the mind

- Min
- Al
- Ch
  -
- Sp
  -
  -
- Sol
  - "Crossword Problem"
  - Proof verification

Crossword Pr

Q. "Woodstoc

A. Don't know

Q. "Woodstoc

A. Joan Baez

Real questior

?

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 $\mathbf{E}(g)$ the average value of g on C^n. Now 1 and 2 are clearly equivalent to the following:

1′. For any boolean function g, $\mathbf{E}(g)\neq 0$ implies $\exists x: s(g,x)\leqslant n-h(n)$.

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 $I\subset[n]$, $\hat{g}(I)=\hat{f}([n]-I)$, therefore $\mathbf{E}(g)=\hat{g}(\varnothing)=\hat{f}([n])$, 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.

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

$2'\to 1'$. 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 $\hat{f}([n])=\hat{g}(\varnothing)=\mathbf{E}(g)=0$, contradicting the premise. ∎

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 !!