What should I talk about?

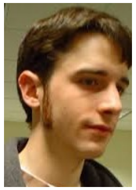
Aspects of Human Communication

Madhu Sudan
Harvard University

Based on many joint works …

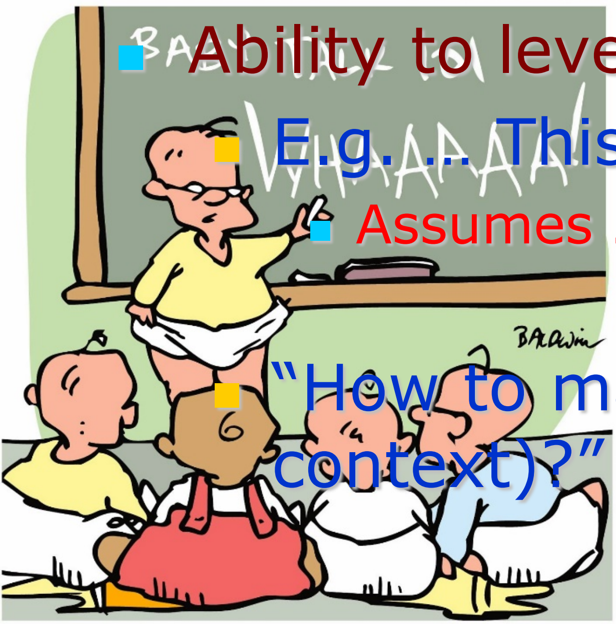
Ingredients in Human Communication

- Ability to start with (nearly) zero context and "learning to communicate by communicating".
  - Children learn to speak … (not by taking courses)
  - Focus of works with Juba and Goldreich
  - "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)

© Mike Baldwin / Cornered



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

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.
- Challenge: Formalize?
  - Work so far … some (toy?) settings

Underlying Model: Communication Complexity

The model (with shared randomness)

x

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

y

$R = \$\$\$$

Alic

Usually studied for lower bounds.
This talk: CC as +ve model.

Bob

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

$f(x,y)$
w.p. 2/3

Aside: Easy CC Problems [Ghazi,Kamath,S'15]

∃ Problems with large inputs and small communication?

- Equality testing:
  - $EQ(x,y) = 1 \Leftrightarrow x = y$; C
- Hamming distance:
  - $H_k(x,y) = 1 \Leftrightarrow \Delta(x,y) \leq k$;
- Small set intersection:
  - $\cap_k (x,y) = 1 \Leftrightarrow$ $\mathrm{wt}(x), \mathrm{wt}(y$
  - $CC(\cap_k) = O(k)$ [Håstad Wi

Protocol:

Fix ECC E $\{0,1\}^n \to \{0,1\}^N$

$poly(k)$ Proto

Use common

to hash $[n] \to$

$x = (x_1, \ldots, x_n)$

$y = (y_1, \ldots, y_n)$

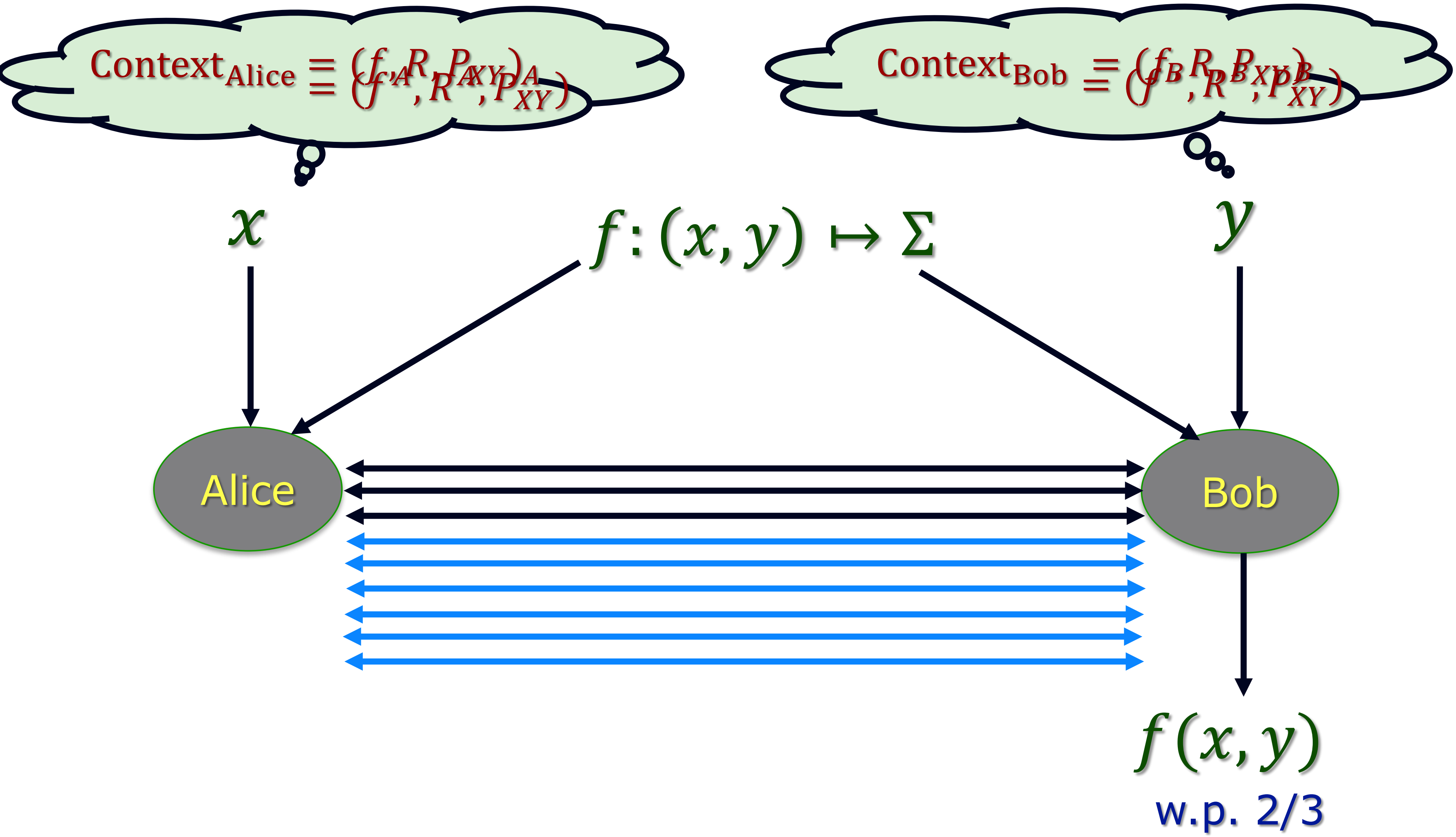
$\langle x, y \rangle \triangleq \sum_i x_i y_i$

Unstated philosophical contribution of CC a la Yao:

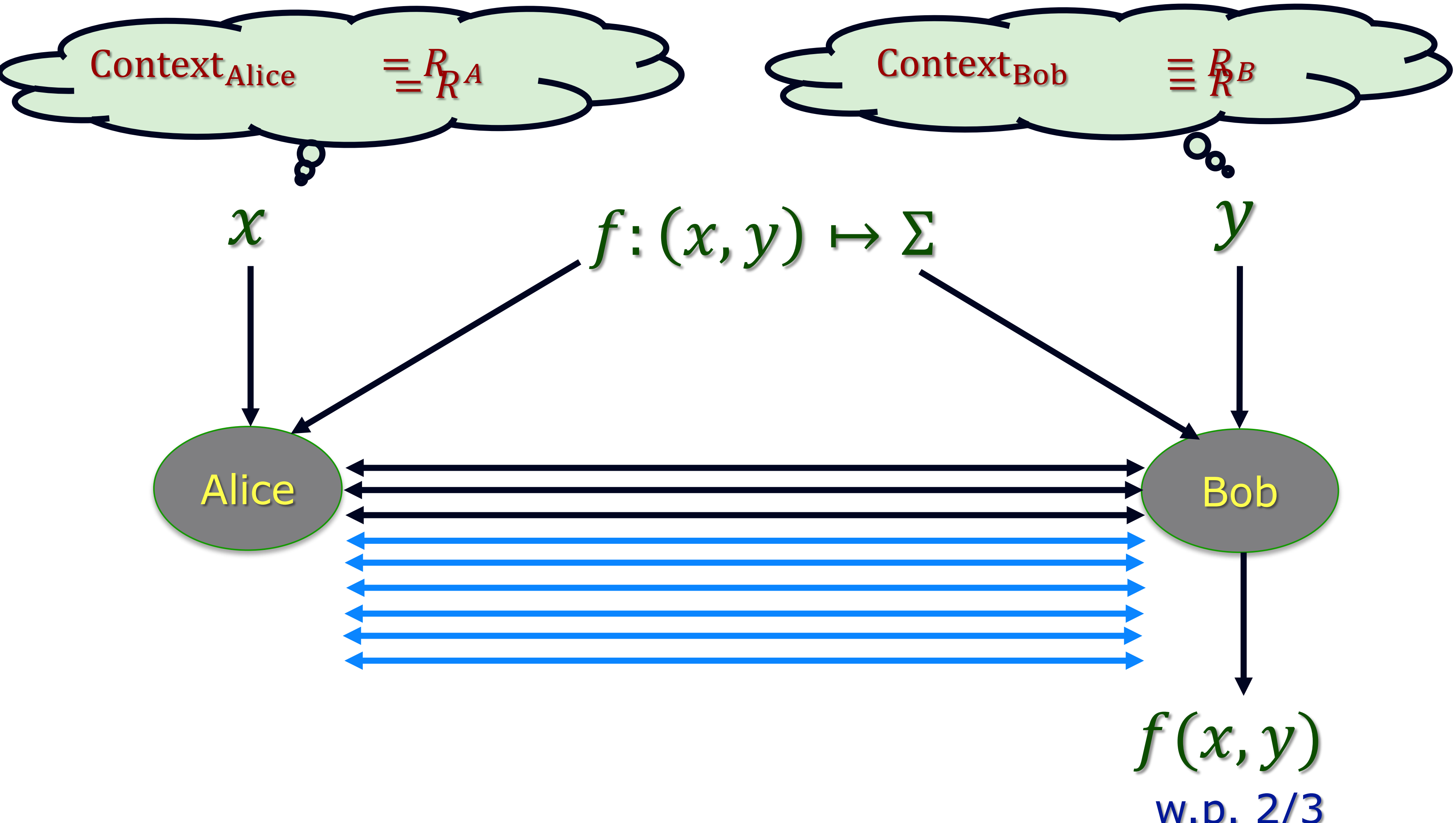
Communication with a focus ("only need to determine $f(x,y)$")

can be more effective (shorter than $|x|, H(x), H(y), I(x;y) \ldots$)

Communication & (Uncertain) Context



1. Imperfectly Shared Randomness

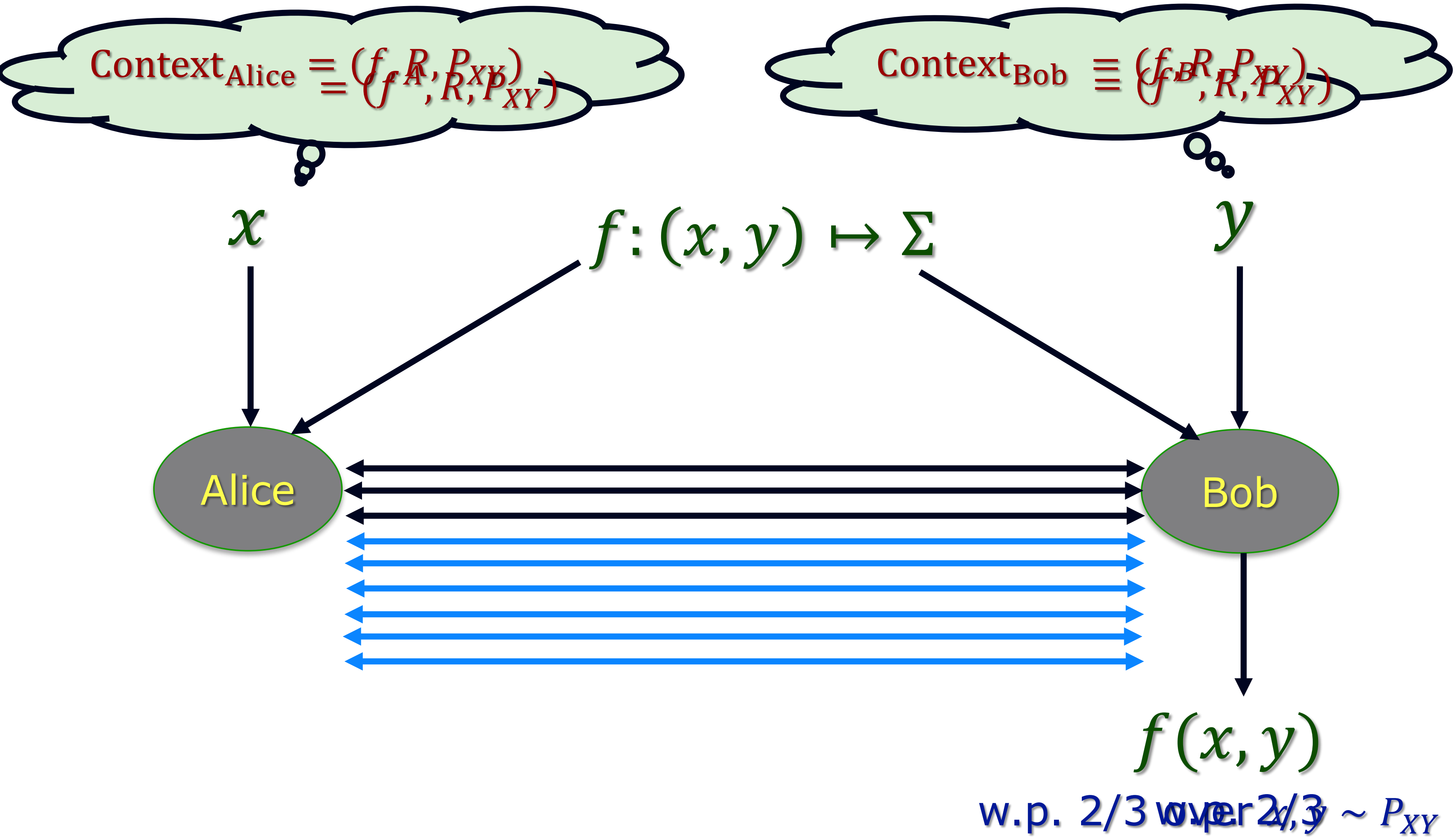


Imperfectly Shared Randomness (ISR)

- Model: $R^A \sim N_\rho(R^B)$ (ρ-correlated iid on each coord.)
- Thm [Bavarian-Gavinsky-Ito'15]: Equality testing has $O(1)$-comm. comp. with ISR.
- Thm [Canonne-Guruswami-Meka-S.'16]: If f has CC k with perfect rand., then it has ISR-CC $O_\rho(2^k)$
- Thm [CGMS] This is tight (for promise problems).

- Complete problem: Estimate $\langle x, y \rangle$ for $x, y \in \mathbb{R}^n$
  - ϵ-approximation needs $\Theta(\epsilon^{-2})$ communication
- Hard problem: Sparse inner product – where x is non-zero only ϵ-fraction of the times.

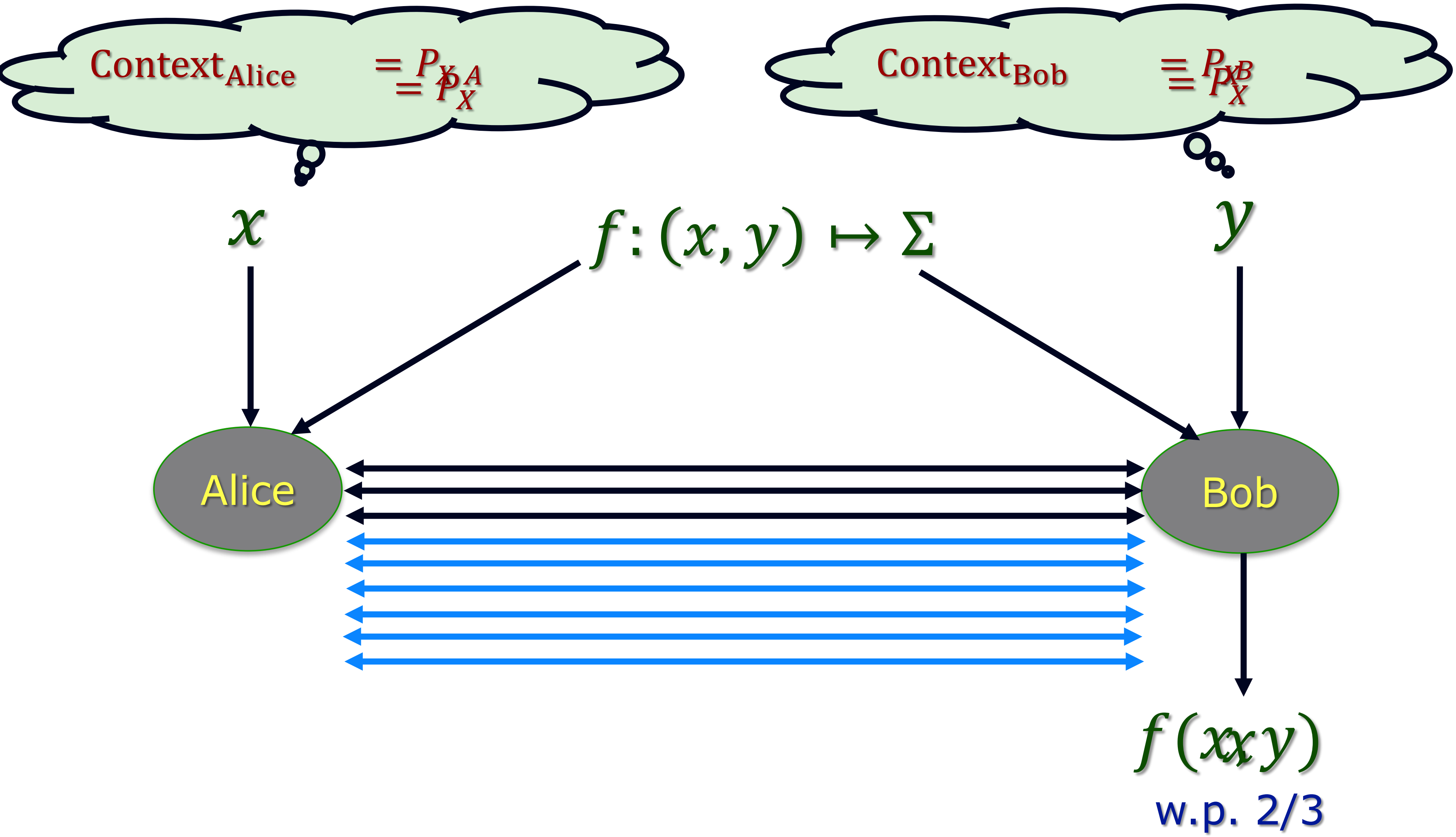
2. Uncertain Functionality



Definitions and results

- Defining problem is non-trivial:
  - Alice/Bob may not "know" f but protocol might!
  - Prevent this by considering entire class of function pairs $\mathcal{G} = \{(f^A, f^B)\}$ that are admissible.
  - Complexity = complexity of $\mathcal{G}$!
- Theorem[Ghazi,Komargodski,Kothari,S. 16]:
  - If P_{XY} arbitrary then there exists $\mathcal{G}$ s.t. every $f \in \mathcal{G}$ has cc = 1, but uncertain-cc$(\mathcal{G}) = \Omega(\sqrt{n})$
  - If P_{XY} = uniform and every $f \in \mathcal{G}$ has one-way cc k, then uncertain-cc$(\mathcal{G}) = O(k)$.
- Theorem[Ghazi-S.,18]: Above needs perfect shared randomness.

3. Compression & (Uncertain) Context



3. (Uncertain) Compression

- Without context: CC = $\log|\Omega|$ (where $x \in \Omega$)
- With shared context, Expected-CC = $H(P_X)$
- With imperfectly shared context, but with shared randomness, Expected-CC = $H\left(P_X^A\right) + \Theta(\Delta)$
  - Where $\Delta = \max_x \left\{\max\left\{\log \frac{P^A(x)}{P^B(x)}, \log \frac{P^B(x)}{P^A(x)}\right\}\right\}$ [JKKS'11]
- Without shared randomness … exact status unknown! Best upper bound ([Haramaty,S'14]):
  - Expected-CC = $O\left(H\left(P_X^A\right) + \Delta + \log\log\Omega\right)$

Compression as a proxy for language

- Information theoretic study of language?
- Goal of language: Effective means of expressing information/action.
- Implicit objective of language: Make frequent messages short. Compression!
- Frequency = Known globally? Learned locally?
  - If latter – every one can't possibly agree on it;
  - Yet need to agree on language (mostly)!
  - Similar to problem of Uncertain Compression.
  - Studied formally in [Ghazi,Haramaty,Kamath,S. ITCS 17]

Part II: Proofs

Well understood … (Goldwasser-Micali-Rackoff, …)

- 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

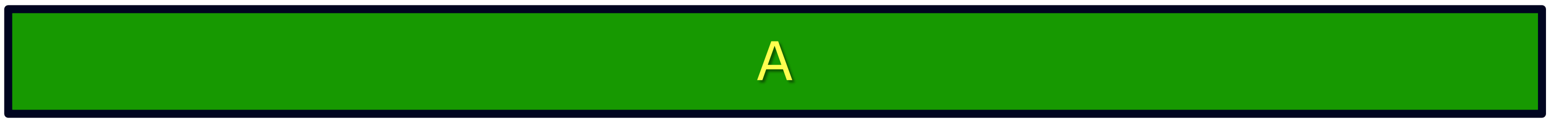
- Small (Constant) Number of Axioms
  - $X \rightarrow Y, Y \rightarrow Z \Rightarrow \; X \rightarrow Z$, Peano, etc.
- Medium Sized Theorem:
  - $\forall x, y, z, n \in \mathbb{N}, \quad x^n + y^n = z^n \rightarrow n \leq 2$ …
- Big Proof:
  - Blah blah blah blah blah blah bla blah blah blah blah blah 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. $\delta(A_P, A_V) \le \delta,\ V^{A_V}(T,\Pi) = 1$
  - 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?

Beyond oracle access: Modelling the mind

- Brain (of verifier) does not just store and retrieve axioms.
- Can make logical deductions too! But should do so feasibly.
- Semi-formally:
  - Let A^t denote the set of axioms "known" to the verifier given t query-proc. time
  - Want $|A^{2t}| \gg |A^t|$, but storage space $\sim |A^0|$
- What is a computational model of the brain that allows this?
  - Cell-probe – No^*. ConsciousTM – Maybe. Etc…

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!