Universal Semantic Communication

Madhu Sudan
MIT CSAIL

Joint work with Brendan Juba (MIT CSAIL).
The Meaning of Bits

- Is this perfect communication?

- What if Alice is trying to send instructions?
  - In other words ... an algorithm
  - Does Bob understand the correct algorithm?
  - What if Alice and Bob speak in different (programming) languages?
Motivation: Better Computing

- Networked computers use common languages:
  - Interaction between computers (getting your computer onto internet).
  - Interaction between pieces of software.
  - Interaction between software, data and devices.

- Getting two computing environments to “talk” to each other is getting problematic:
  - time consuming, unreliable, insecure.

- Can we communicate more like humans do?
Some modelling

Say, Alice and Bob know different programming languages. Alice wishes to send an algorithm A to Bob.

Bad News: Can’t be done
- For every Bob, there exist algorithms A and A’, and Alices, Alice and Alice’, such that Alice sending A is indistinguishable (to Bob) from Alice’ sending A’

Good News: Need not be done.
- From Bob’s perspective, if A and A’ are indistinguishable, then they are equally useful to him.

What should be communicated? Why?
Aside: Why communicate?

- Classical “Theory of Computing”

- Issues: Time/Space on DFA? Turing machines?

- Modern theory:

- Issues: Reliability, Security, Privacy, Agreement?

- If communication is so problematic, then why not “Not do it”?
(Selfish) Motivations for Communication

- Bob speaks to some environment (a collection of entities).

- Why? Has some goal!
  - “Control”: Wants to alter the state of the environment.
  - “Intellectual”: Wants to glean knowledge (about universe/environment).

- Claim: By studying the goals, can enable Bob to overcome linguistic differences (and achieve goal).
Rest of the talk

- **Part I:** Bob is computationally limited but wishes to solve hard problem, and Alice can solve the problem.

- **Part II:** Bob is a teacher and wants to test student’s ability.

- **Part III:** Generic goals.
Part I: A Computational Goal
Modelling the communicator (Bob)

- Bob: $\Omega \times \Sigma^k \rightarrow \Omega \times \Gamma^\ell$, where $\Omega =$ countable state space, $\Sigma^k =$ input signals, $\Gamma^\ell =$ output signals.

- Alice similar
Computational Goal for Bob

- Bob is probably polynomial time bounded. Wants to decide membership in set $S$.

- Alice is computationally unbounded, does not speak the same language as Bob, but is “helpful”.

- What kind of sets $S$?
  - E.g., undecidable?, decidable? PSPACE, NP, BPP?
Setup

Bob

\[ x \in S? \]

\[ R \leftarrow $$$ \]

Different from IP:
In IP Bob does not trust Alice, while here he does not understand her.

\[ f(x, R, a_1, \ldots, a_k) = 1? \]

\[ q_k \]

\[ a_k \]

Hopefully \[ x \in S \iff f(\cdots) = 1 \]
Helpful Alice?

For Bob to have a non-trivial interaction, Alice must be:

- Intelligent: Capable of deciding if $x$ in $S$.
- Cooperative: Must communicate this to Bob.

Formally:

Alice is $S$-helpful if $\exists$ probabilistic poly time (ppt) Bob $B'$ s.t.

$\forall$ initial state of mind $\sigma$,

$A(\sigma) \leftrightarrow B'(x)$ accept w.h.p. iff $x \in S$. 
Successful universal communication

- Bob should be able to talk to any S-helpful Alice and decide S.

- Formally,

\[
Ppt \ B \text{ is } S\text{-universal} \text{ if for every } x \in \{0, 1\}^* \\
\quad A \text{ is } S\text{-helpful } \Rightarrow [A \leftrightarrow B(x)] = 1 \text{ iff } x \in S \text{ (whp)}.
\]

\[
A \text{ is not } S\text{-helpful } \Rightarrow \text{Nothing!!}
\]

Or should it be ...

\[
A \text{ is not } S\text{-helpful } \Rightarrow [A \leftrightarrow B(x)] = 1 \text{ implies } x \in S.
\]
Main Theorem

- If $S$ is PSPACE-complete, then there exists a $S$-universal Bob (generalizes to other checkable sets $S$).

- Conversely, if there exists a $S$-universal Bob, then $S$ is in PSPACE.

- In other words:
  - If $S$ is moderately stronger than what Bob can do on his own, then attempting to solve $S$ leads to non-trivial (useful) conversation.
  - If $S$ too strong, then leads to ambiguity.
  - Uses $IP=PSPACE$ [LFKN, Shamir]
Few words about the proof

- **Positive result:** Enumeration + Interactive Proofs
  
  Guess: Interpreter; $x \in S$?

Proof works $\Rightarrow x \in S$; Doesn't work $\Rightarrow$ Guess wrong.

Alice $S$-helpful $\Rightarrow$ Interpreter exists!
Proof of Negative Result

- \( L \) not in PSPACE implies \( \text{Bob makes mistakes} \).
  - Suppose Alice answers every question so as to minimize the conversation length.
    - (Reasonable effect of misunderstanding).
  - Conversation comes to end quickly.
  - Bob has to decide.
  - Conversation + Decision simulatable in PSPACE (since Alice’s strategy can be computed in PSPACE).
  - Bob must be wrong if \( S \) is not in PSPACE.
  - Warning: Only leads to finitely many mistakes.
Part II: Generic Goals
Generically

- Bob interacts with an environment (collection of Alice(s)).

- What should goal depend on?
  - States of Bob? Then how can Bob adapt to Alice?
  - State of Alice(s)? Bob doesn’t know this!
  - Transcript of interaction? Does this mean the same thing for different Alice/Bob pairs?
An Analogy: Multiparty Computation

- Need to model generic multiparty computation, to present general protocols for “secure, private, multiparty computation”.

- Modelled by “Ideal Trusted Party”

$\{(f_a, f_b, f_c, f_d, f_e)\}$
Generic Goals

- **Framework:** Bob talks to Alice thru Interpreter

- **Roles:**
  - Bob defines the **Goal** (though his actions may depend also on what the interpreter hears from Alice).
  - Alice comes from class $\tilde{A}$; Interpreter from $\tilde{I}$
  - Alice is helpful if Bob achieves his goal with her thru some Interpreter in $\tilde{I}$
  - Interpreter is universal if Bob achieve his goal for every helpful Alice in $\tilde{A}$. 

Generic Helpfulness, Universality

Consider: Class of Alices $\mathcal{A}$, Class of Interpreters $\mathcal{I}$ and some goal given by Bob $B$

- $(B, \mathcal{I})$-Helpful: Alice helpful to Bob via some Interpreter in $\mathcal{I}$.
- $(B, \mathcal{A})$-Universal: Interpreter works with all Alice in $\mathcal{A}$.

Theorem: “Forgiving”, “verifiable” Goals can be achieved universally.

- “Forgiving” – no finite prefix of interaction should rule out achievement of Goal.
- “Verifiability” ...
Typical Goals

- Intent of Goals: Usually depend on state of Alice!
- Realizable goals: Can only depend on state of Bob, Interpreter and interactions.
- Translating Intent to Realizable Goal: non-trivial.
Part III: Intellectual Curiosity
Setting: Bob more powerful than Alice

- What should Bob’s Goal be?
  - Can’t use Alice to solve problems that are hard for him.
  - Can pose problems and see if she can solve them. E.g., Teacher-student interactions.
  - But how does he verify “non-triviality”?
  - What is “non-trivial”? Must distinguish …
Setting: Bob more powerful than Alice

- Concretely:
  - Bob capable of $\text{TIME}(n^{10})$.
  - Alice capable of $\text{TIME}(n^3)$ or nothing.
  - Can Bob distinguish the two settings?

- Definition:
  Alice is $n^{3-\epsilon}$-helpful
  
  if $\exists$ Bob $B' \in \text{TIME}(n^{3-\epsilon})$ s.t.
  
  $\forall S \in \text{TIME}(n^3)$, and $\forall$ initial state of mind $\sigma$, $\sigma$
  
  $A(\sigma) \leftrightarrow B'(x_1, \ldots, x_n)$ computes $S(x_1), \ldots, S(x_n)$.

- Theorem: There exists a universal Bob that distinguishes helpful Alices from trivial ones.

- Moral: Language (translation) should be simpler than problems being discussed.
Conclusions

- Communication of “meaning/context” is feasible; provided goals are explicit.

- Verifying “goal achievement” for non-trivial goals is the (only?) way to learn languages.

- Currently the learning is slow ... is this inherent?
  - Better class of Alices?

- What are interesting goals, and how can they be verified?
Thank You!
Computers Communicate!

- Classical “Theory of Computing”

\[ X \xrightarrow{F} F(X) \]

- Issues: Time/Space on DFA? Turing machines?
- Modern theory:

- Issues: Reliability, Security, Privacy, Agreement?
Computers Communicate! How? Why?

- Classical Introduction to Theory of Computing

- **Bad News:** Can’t be done
  - For every Bob, there exist algorithms A and A’, and Alices, Alice and Alice’, such that the two are indistinguishable to Bob.

- **Good News:** Need not be done.
  - From Bob’s perspective, if A and A’ are indistinguishable, then they are equally useful to him.

- What should be communicated? Why?
Computers Communicate!

- **Classical Introduction to Theory of Computing**

- **Bad News:** Can’t be done
  - For every Bob, there exist algorithms $A$ and $A'$, and Alices, Alice and Alice', such that the two are indistinguishable to Bob.

- **Good News:** Need not be done.
  - From Bob’s perspective, if $A$ and $A'$ are indistinguishable, then they are equally useful to him.

- What should be communicated? Why?