

Universal *Semantic* Communication

Madhu Sudan
MIT CSAIL

Joint work with Brendan Juba (MIT).

An fantasy setting (SETI)



Alice

010010101010001111001000

No common language!
Is meaningful
communication possible?



Bob

What should Bob's response be?

If there are further messages, are they reacting to him?

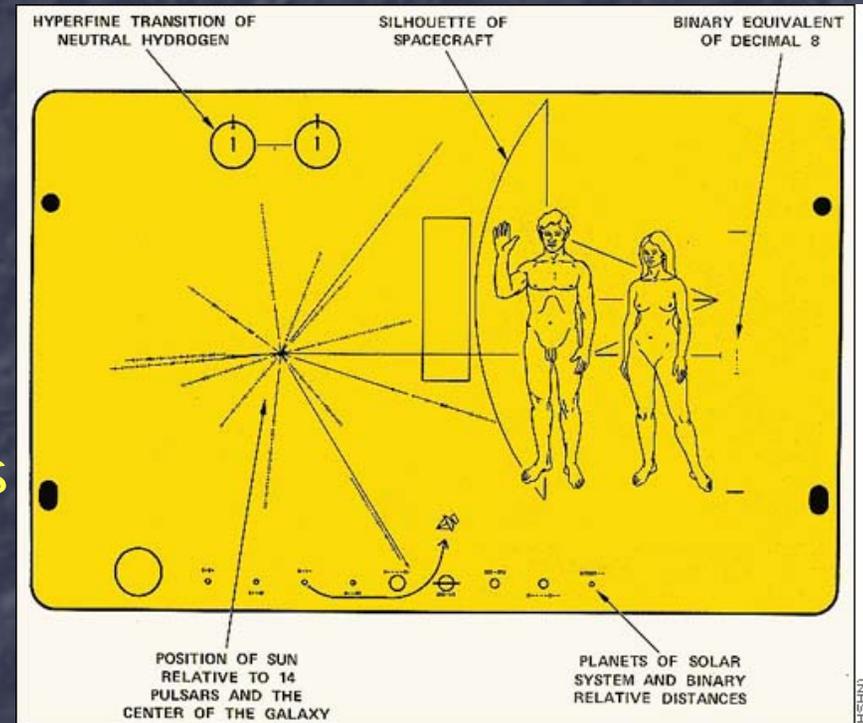
Is there an intelligent Alien (Alice) out there?

Pioneer's face plate

Why did they put this image?

What would you put?

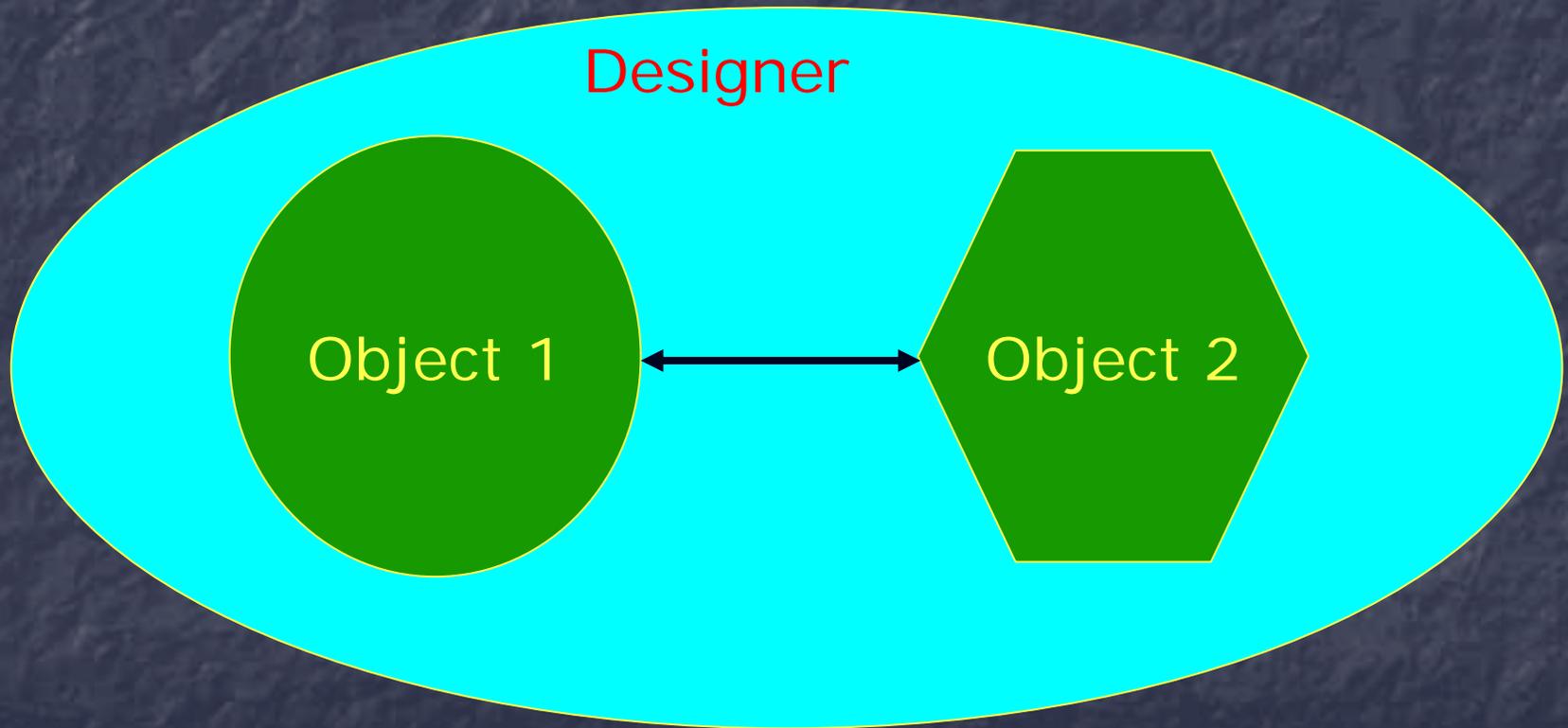
What are the assumptions and implications?



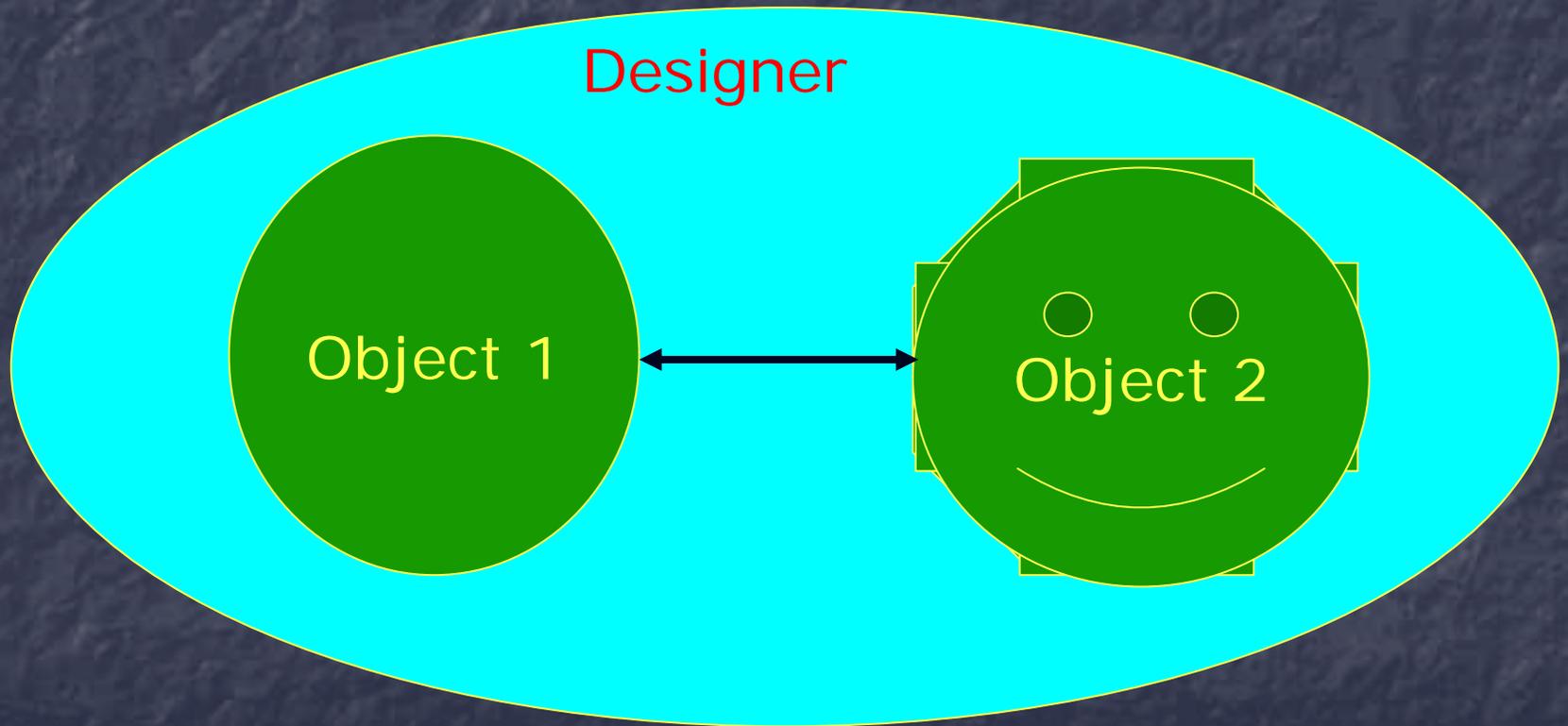
Motivation: Better Computing

- 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?

Classical Paradigm for interaction



New paradigm



Robust interfaces

- Want one interface for all “Object 2”s.
- Can such an interface exist?
- What properties should such an interface exhibit?
- Puts us back in the “Alice and Bob” setting.

Goal of this talk

- Definitional issues and a definition:
 - What is successful communication?
 - What is intelligence? cooperation?
- Theorem: "If Alice and Bob are intelligent and cooperative, then communication is feasible" (in one setting)
- Proof ideas:
 - Suggest:
 - Protocols, Phenomena ...
 - Methods for proving/verifying intelligence

A first attempt at a definition

- Alice and Bob are “universal computers” (aka programming languages)
- Have no idea what the other’s language is!
- Can they learn each other’s language?

- **Good News:** Language learning is finite. Can enumerate to find translator.

- **Bad News:** No third party to give finite string!
 - Enumerate? Can’t tell **right/wrong** ☹

Communication & Goals

- Indistinguishability of Right/Wrong: Consequence of "communication without goal".
- Communication (with/without common language) ought to have a "Goal".
- Before we ask how to improve communication, we should ask why we communicate?

*"Communication is not an end in itself,
but a means to achieving a Goal"*

Part I: A Computational Goal

Computational Goal for Bob

- Bob wants to solve hard computational problem:
 - Decide membership in set S .
- Can Alice help him?
- What kind of sets S ? E.g.,
 - $S = \{\text{set of programs } P \text{ that are not viruses}\}$.
 - $S = \{\text{non-spam email}\}$
 - $S = \{\text{winning configurations in Chess}\}$
 - $S = \{(A,B) \mid A \text{ has a factor less than } B\}$

Review of Complexity Classes

- **P (BPP)** – Solvable in (randomized) polynomial time (Bob can solve this without Alice's help).
- **NP** – Problems where solutions can be verified in polynomial time (contains factoring).
- **PSPACE** – Problems solvable in polynomial space (quite infeasible for Bob to solve on his own).
- **Computable** – Problems solvable in finite time. (Includes all the above.)
- **Uncomputable** (Virus detection. Spam filtering.)

Which problems can you solve
with communication?

Setup

Which class
of sets?

Bob

$x \in S?$

$R \leftarrow \text{\$}\text{\$}\text{\$}$

Alice

q_1

a_1

•

•

•

q_k

a_k

$f(x, R, a_1, \dots, a_k) = 1?$

Hopefully $x \in S \Leftrightarrow f(\dots) = 1$

Intelligence & Cooperation?

- For Bob to have a non-trivial exchange 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.

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

(independent of the history)

Successful universal communication

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

- Formally,

Ppt B is S -universal if for every $x \in \{0, 1\}^*$

– $x \in S$ and A is S -helpful $\Rightarrow (A \leftrightarrow B(x)) = 1$ (whp).

– $(A \leftrightarrow B(x)) = 1$ whp (and A is S -helpful) $\Rightarrow x \in S$.

Main Theorem

- - If S is PSPACE-complete (aka Chess),
then there exists an S -universal Bob.
(Generalizes to many other problems in PSPACE.)
- - If there exists an S -universal Bob
then S is in PSPACE.
- In English:
 - 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]

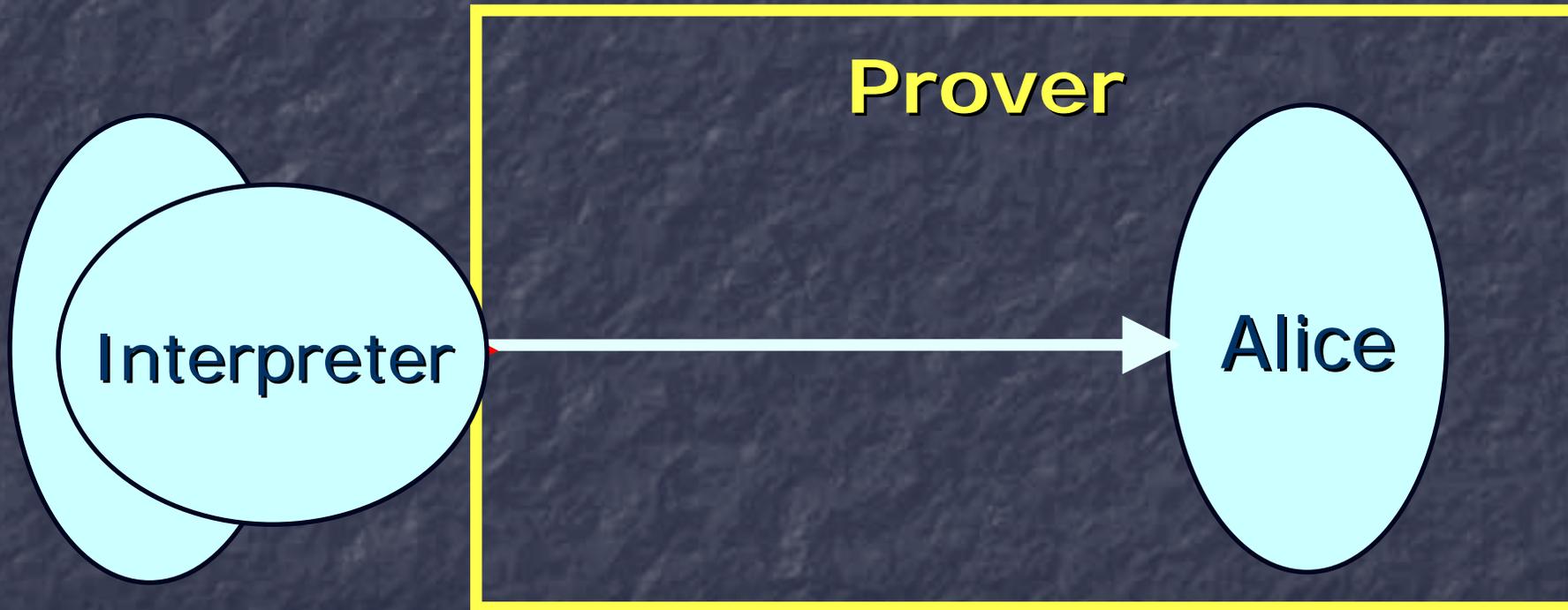
Contrast with Interactive Proofs

- **Similarity:** Interaction between Alice and Bob.
- **Difference:** Bob does not **trust** Alice.
(In our case Bob does not **understand** Alice).
- **Famed (hard) theorem:** $IP = PSPACE$.
 - Membership in PSPACE solvable S can be proved interactively to a probabilistic Bob.
 - Needs a PSPACE-complete prover Alice.

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!

Few words about the proof

- Positive result: Enumeration + Interactive Proofs
- Negative result:
 - Suppose Alice answers every question so as to minimize the conversation length.
 - Reasonable(?) misunderstanding.
 - Conversation comes to end quickly.
 - Bob has to decide.
 - Decision can be computed in PSPACE (since Alice's strategy can be computed in PSPACE).
 - Bob must be wrong if L is not in PSPACE.
 - **Warning:** Only leads to finitely many mistakes.

Is this language learning?

- End result promises no language learning: Merely that Bob solves his problem.
- In the process, however, Bob learns Interpreter!
- But this may not be the right Interpreter.
- All this is Good!
 - No need to distinguish indistinguishables!

Part II: Other Goals?

Goals of Communication

- Largely unexplored (at least explicitly)!
- Main categories
 - Remote Control:
 - Laptop wants to print on printer!
 - Buy something on Amazon
 - Intellectual Curiosity:
 - Learning/Teaching
 - Listening to music, watching movies
 - Coming to this talk
 - Searching for alien intelligence
 - May (not) involve common background

Extending results to other goals

- **Generic Goal** (for Bob): Boolean function G of
 - Private input, randomness
 - Interaction with Alice through Interpreter
 - Environment (Altered by actions of Alice)
- Should be
 - **Verifiable**: Easily computable function of interaction;
 - **Complete**: Achievable with common language.
 - **Non-trivial**: Not achievable without Alice.

Generic Goals

- Can define Goal-helpful; Goal-universal; and prove existence of Goal-universal Interpreter for all Goals.
- **Claim:** Captures all communication (unless you plan to accept random strings).
- Modelling natural goals is still interesting. E.g.
 - Printer Problem: Bob(x): Alice should say x.
 - Intellectual Curiosity: Bob: Send me a "theorem" I can't prove, and a "proof".
 - Proof of Intelligence (computational power):
Bob: given f , x ; compute $f(x)$.
- **Conclusion:** (Goals of) Communication can be achieved w/o common language

Role of common language?

- If common language is not needed (as we claim), then why do intelligent beings like it?
 - Our belief: To gain efficiency.
 - Reduce # bits of communication
 - # rounds of communication
- Topic for further study:
 - What efficiency measure does language optimize?
 - Is this difference asymptotically significant?

Summary

- Communication should strive to satisfy one's goals.
- If one does this "understanding" follows.
- Can enable understanding by dialog:
 - Laptop -> Printer: Print <file>
 - Printer: But first tell me
 - "If there are three oranges and you take away two, how many will you have?"
 - Laptop: One!
 - Printer: Sorry, we don't understand each other!
 - Laptop: Oh wait, I got it, the answer is "Two".
 - Printer: All right ... printing.

Further work

- Criticism:
 - PSPACE Alice?
 - Exponential time learning (enumerating Interpreters)
 - Necessary in our model.
- What are other goals of communication?
- What are assumptions needed to make language learning efficient?

<http://theory.csail.mit.edu/~madhu/papers/juba.pdf>

Thank You!