Semantic Goal-Oriented Communication

Madhu Sudan
Microsoft Research + MIT

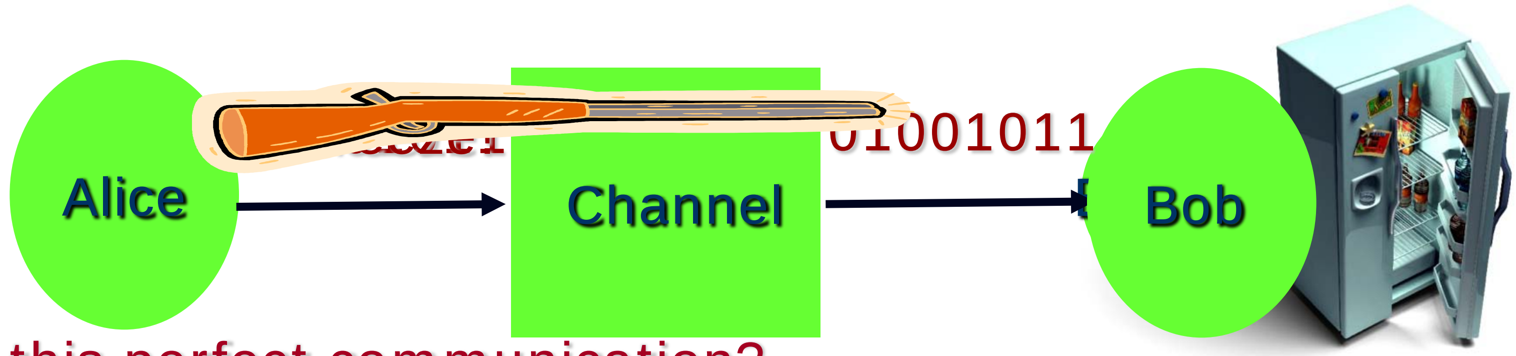
Joint with **Oded Goldreich** (Weizmann) and **Brendan Juba** (MIT).

Disclaimer

- Work in progress (for ever) …

- Comments/Criticisms welcome.

The Meaning of Bits



- Is this perfect communication?

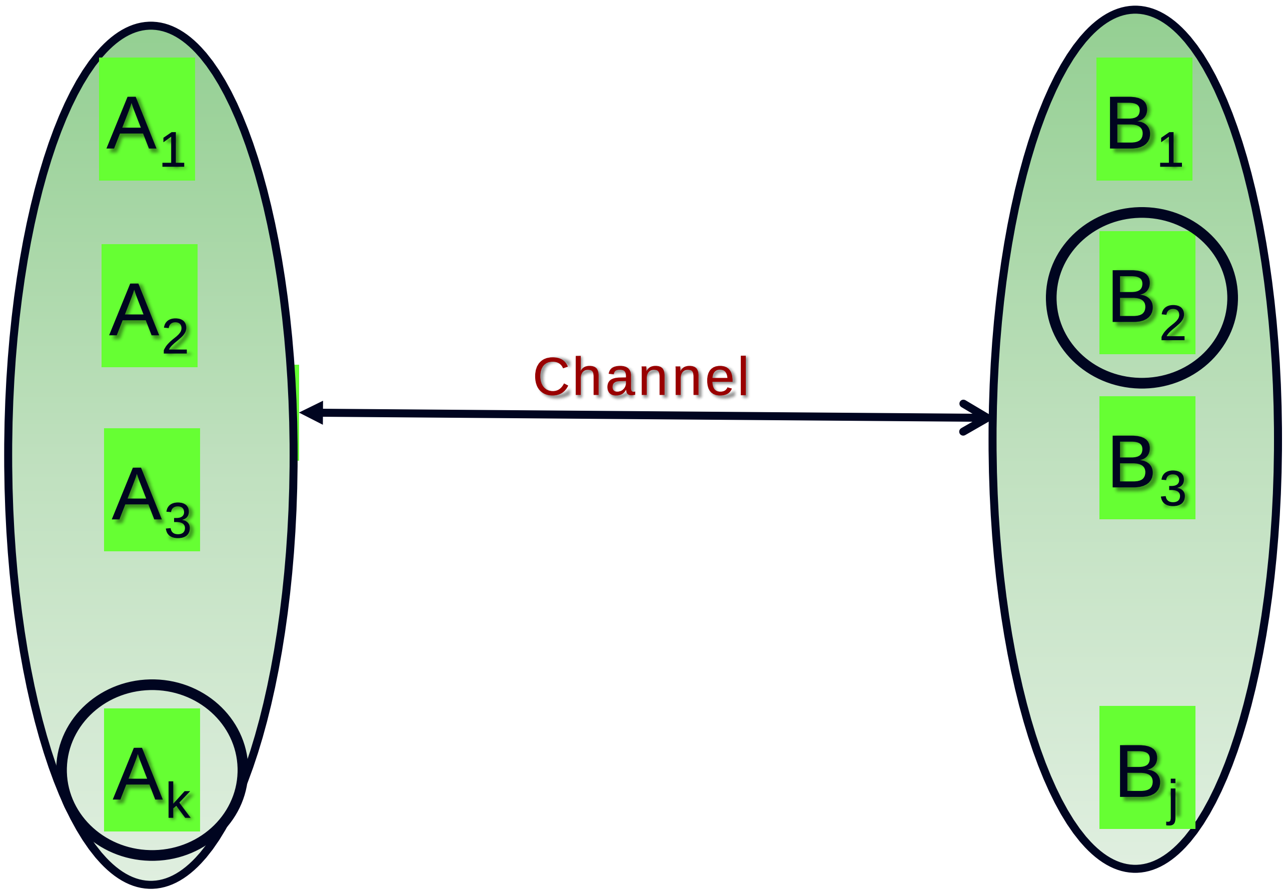
- What if Alice is trying to send instructions?
  - Aka, an algorithm
  - Does Bob understand the correct algorithm?
  - What if Alice and Bob speak in different (programming) languages?

Miscommunication (in practice)

- Exchanging (powerpoint) slides.
  - Don't render identically on different laptops.
- Printing on new printer.
  - User needs to "learn" the new printer, even though printer is quite "intelligent".
- Many such examples …
  - In all cases, sending bits is insufficient.
  - Notion of meaning … intuitively clear.
  - But can it be formalized?
    - Specifically? Generically?
    - While conforming to our intuition

Modelling Miscommunication

Classical Shannon Model
Semantic Communication Model



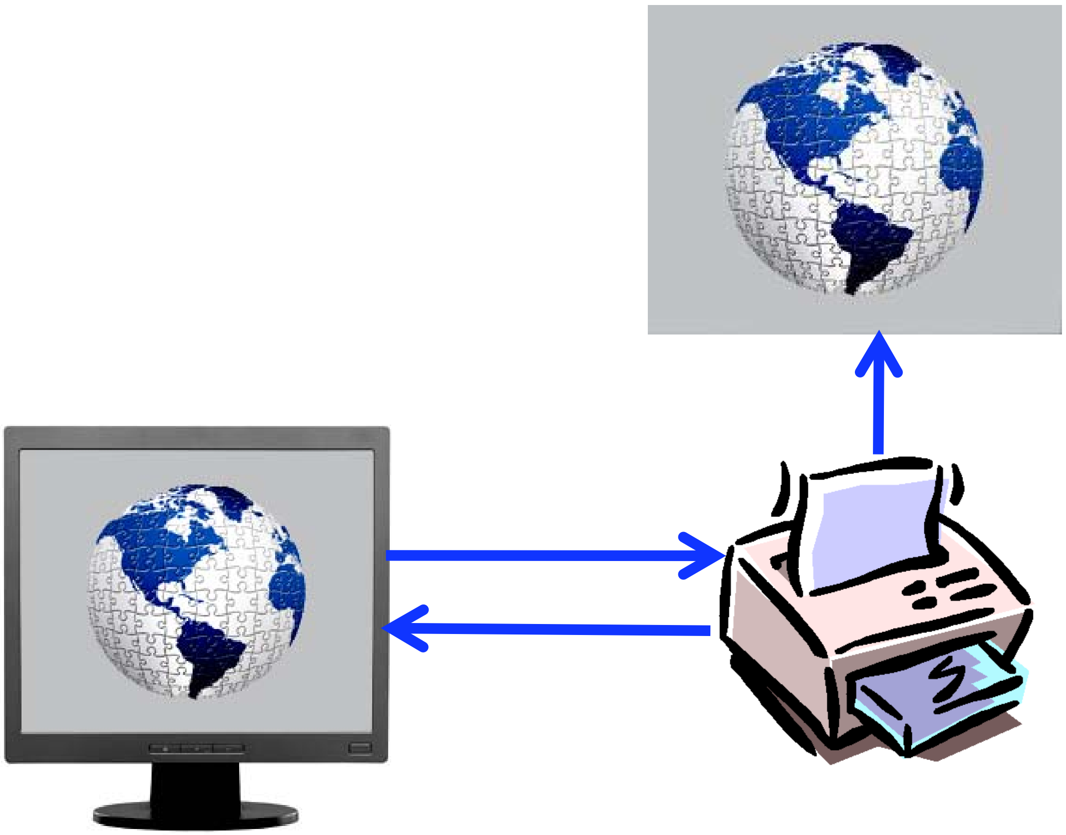
Basic issues

- Source of Miscommunication:
  - A_i doesn't know j
  - B_j doesn't know i
- But what do they wish to achieve?
  - Distinguish B_j from B_k?
    - What if they are indistinguishable?
- Thesis: Communication ought to have Goal!!!
  - Alice/Bob should strive to achieve Goal.
  - Is there a specific Goal to all communication?
  - Are there many possible Goals?
  - Goal specifies problem, but what is a solution?

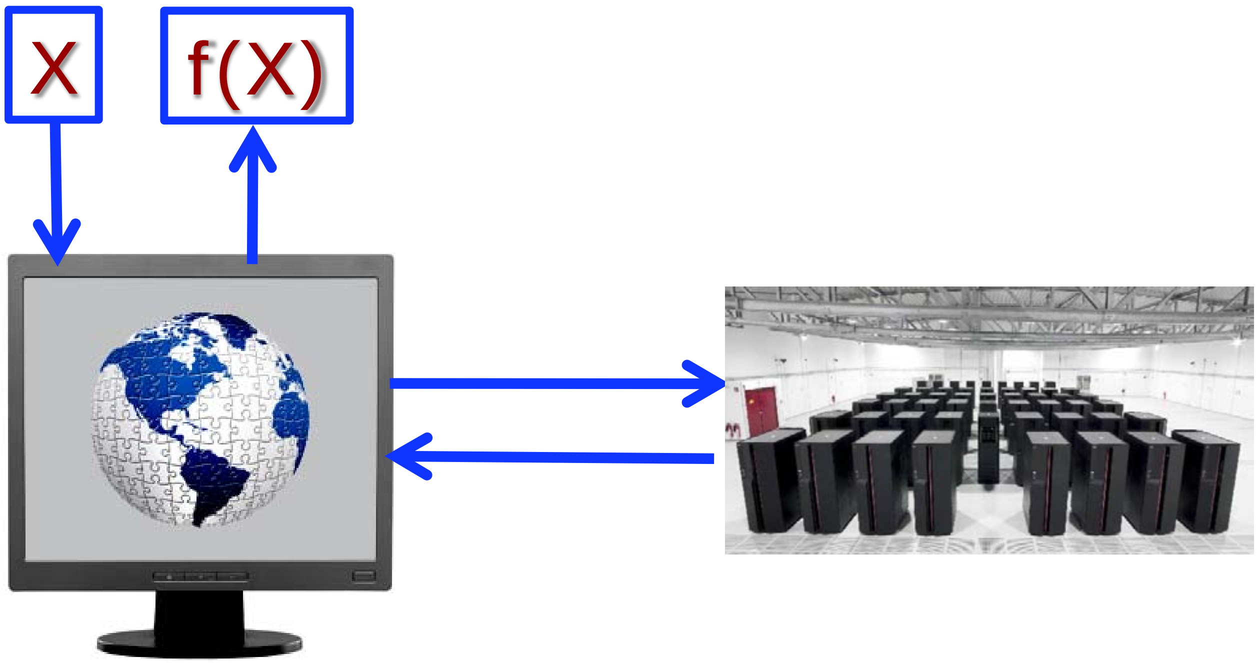
Examples of Goals

- In future slides:
- User  communicates/interacts with Server.
- Will try to look at 's goal.

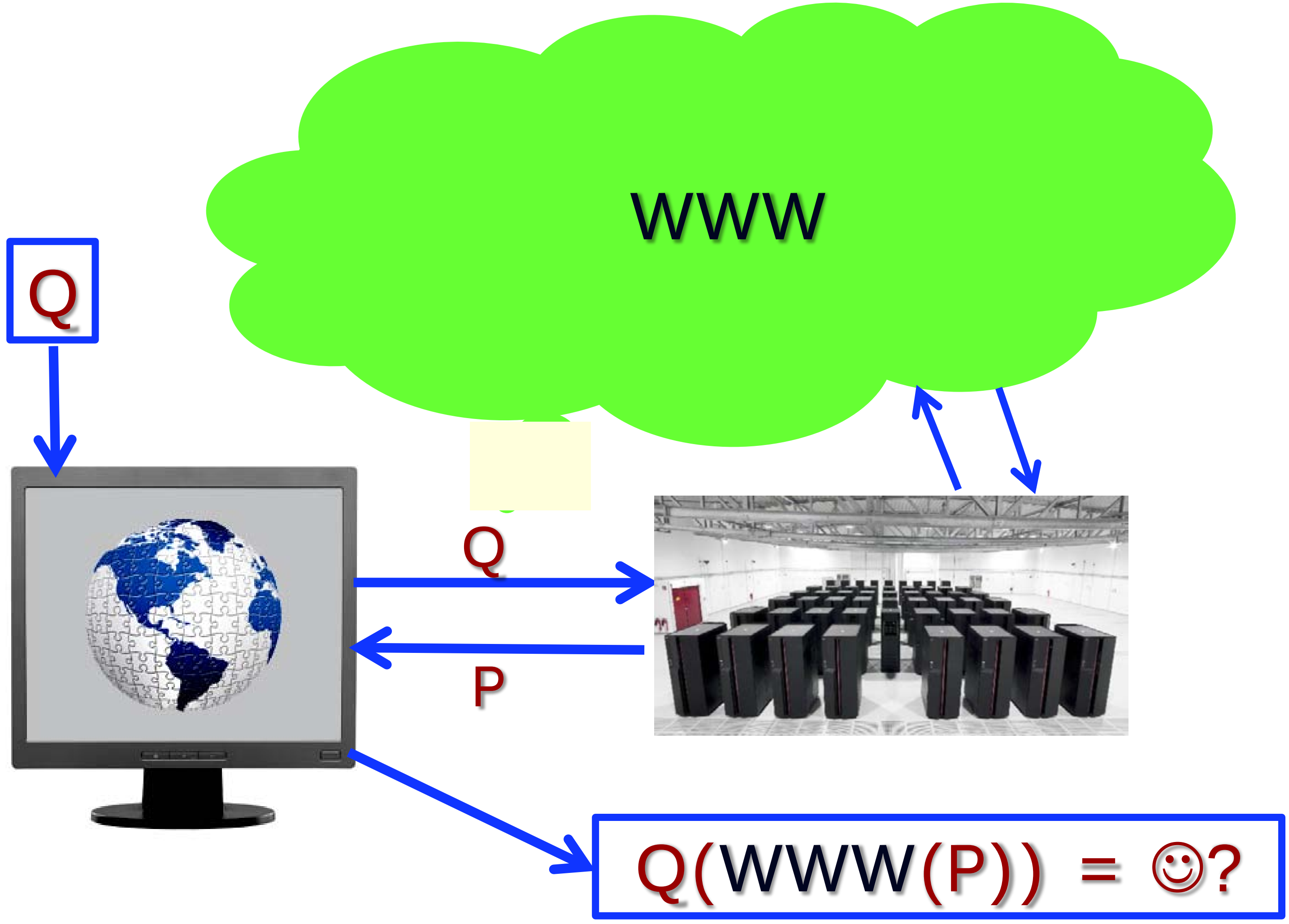
Communication: Example 1 (Printing)



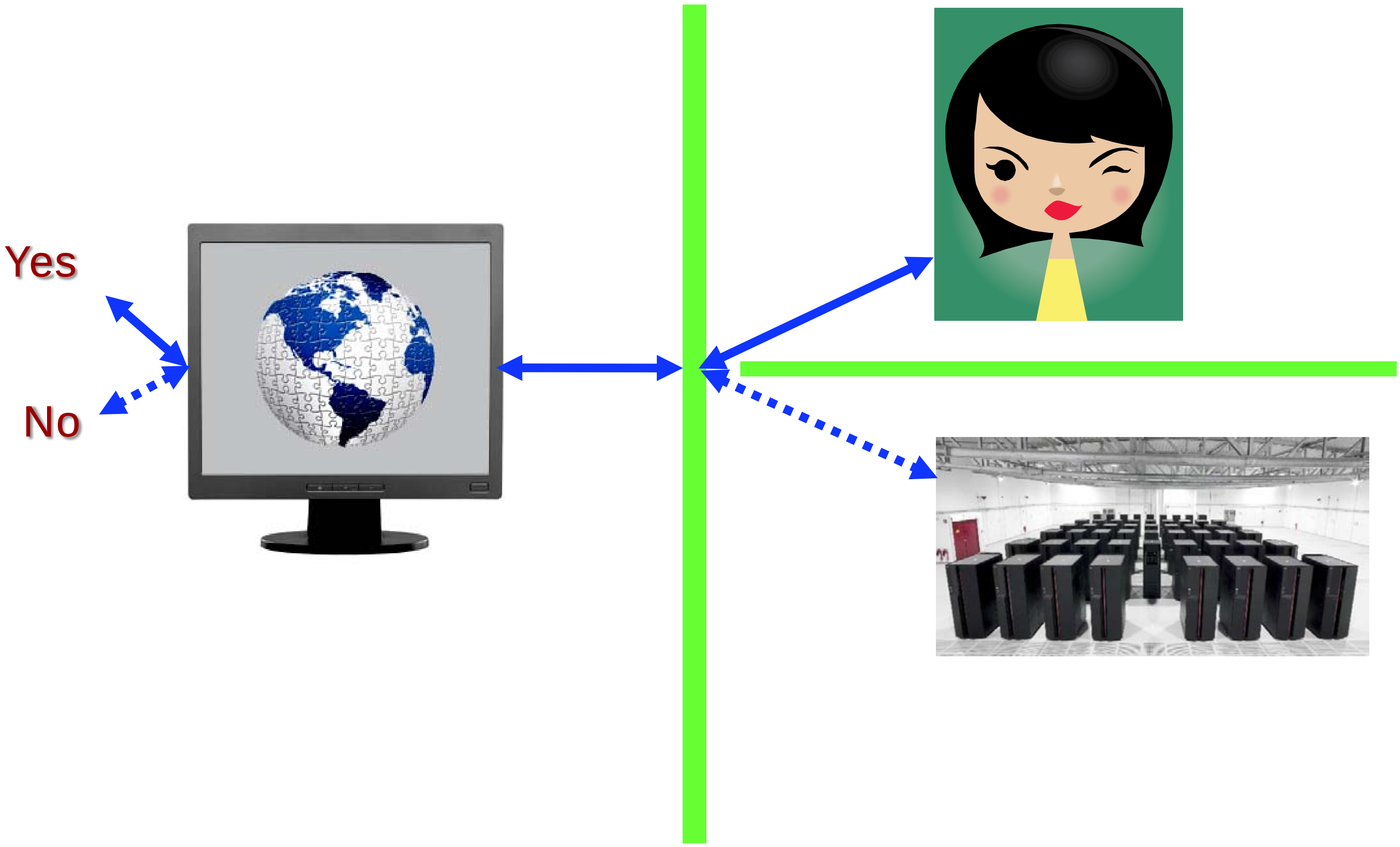
Communication: Ex. 2 (Computation)



Communication: Ex. 3 (Web search)

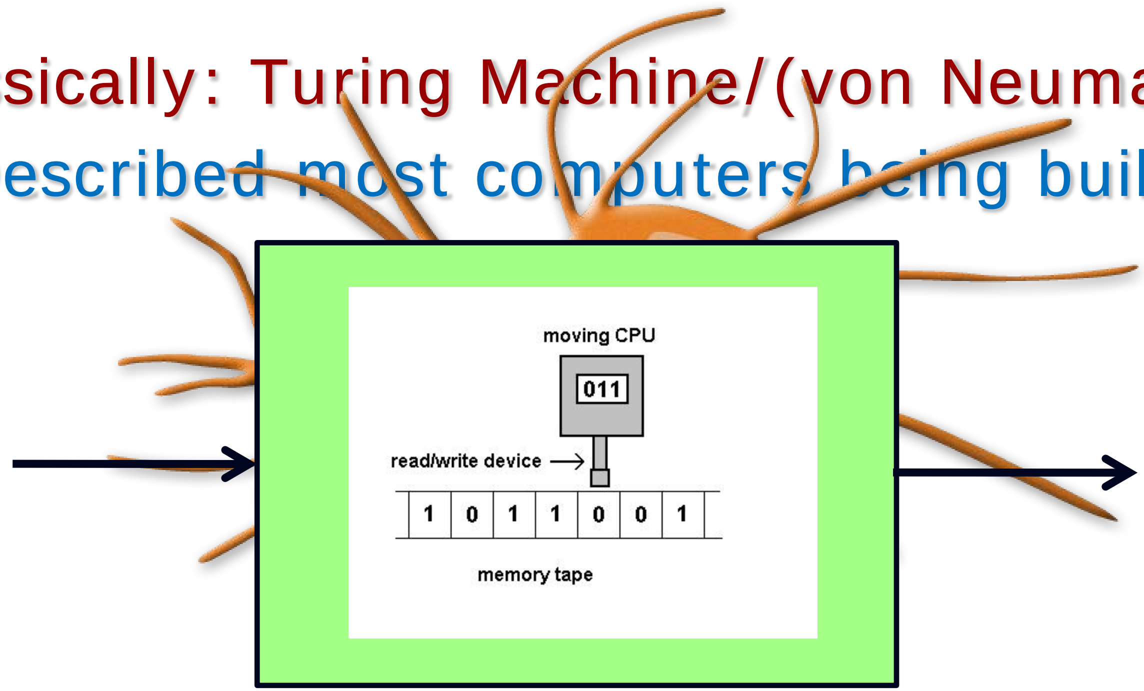


Communication: Ex. 4 (Intelligence?)



Aside: Modelling Computing

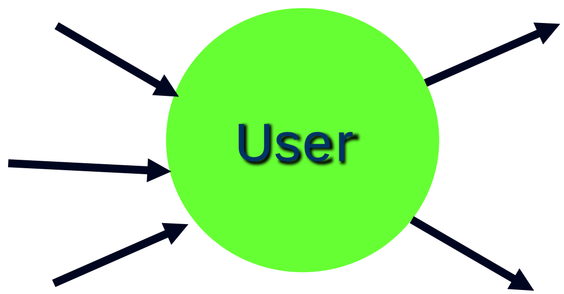
- Classically: Turing Machine/(von Neumann) RAM.
  - Described most computers being built?



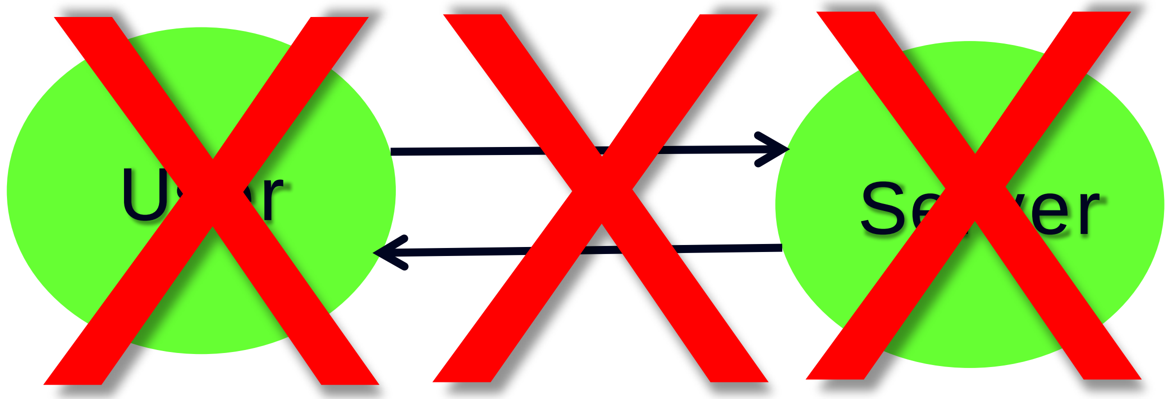
- Modern computers: more into communication than computing.
  - What is the mathematical model of a communicating computer? Why do they communicate? What are all the “communication problems”? What is universality?

Modelling User/Interacting agents

- (standard AI model)
- User has state and input/output wires.
    - Defined by the map from current state and input signals to new state and output signals.



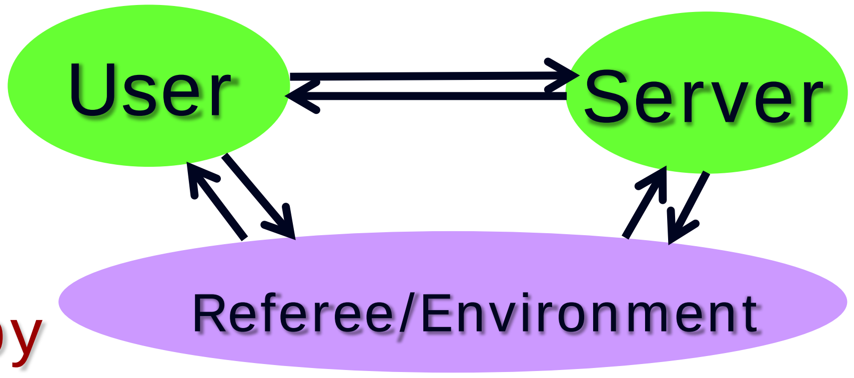
Generic Goal?



- Goal = function of ?
  - User? – But user wishes to change actions to achieve universality!
  - Server? – But server also may change behaviour to be helpful!
  - Transcript of interaction? – How do we account for the many different languages?

Generic Goals

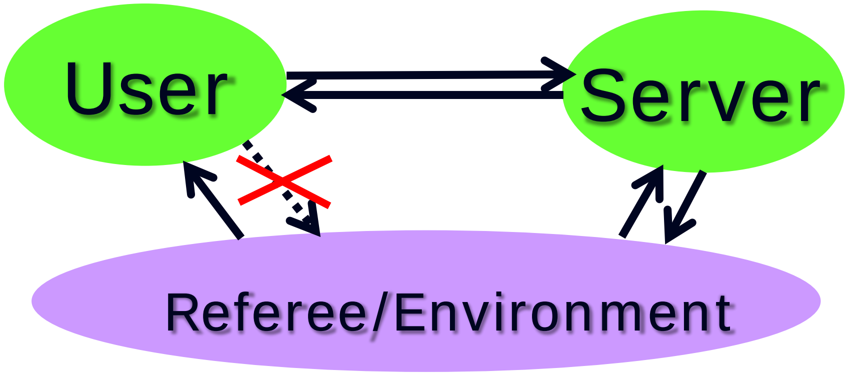
- Key Idea: Introduce 3rd entity: Referee
  - Poses tasks to user.
  - Judges success.



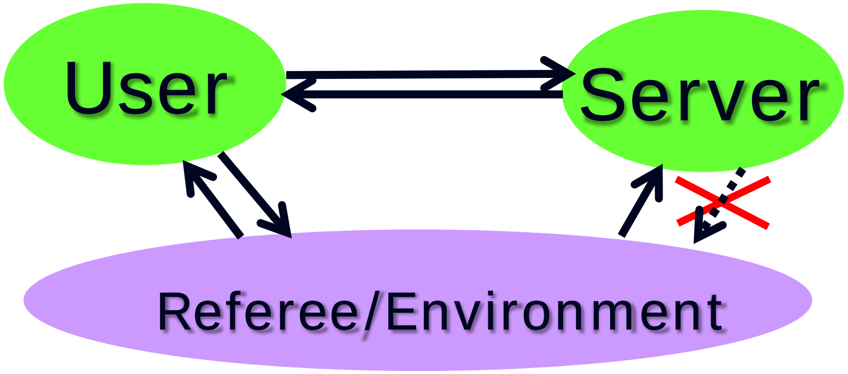
- Generic Goal specified by
  - Referee (just another agent)
  - Boolean Function determining if the state evolution of the referee reflects successful achievement of goal.
  - Class of users/servers.

Generic Goals

- Pure Control



- Pure Informational



Sensing & Universality

- To achieve goal:
  - Server should be "helpful"
  - User should be able to "sense progress".
    - I.e., user should be compute a function that mimics referee's verdict.
- General positive result [GJS '09]:
  - Generic goals (with appropriate definitions) universally achievable if ∃ sensing function.
- General negative result [GJS '09]:
  - Sensing is necessary (in sufficiently general classes of users/servers).

Implications of “Universality”

- Standard question in linguistics, cognition …
  - What is a precondition for two entities to come to some “common language”?
  - Standard answers:
    - Humans seem to need little commonality (a child can learn any language)
    - But humans share enormous common physical needs and have large common genetic code?
    - Is all this necessary?
  - Our Answer: No. Compatible goals suffice.

Concrete Example: Computation

Computational Goal for User

- User wants to compute function f on input x.
- Setting:
  - User is prob. poly time bounded.
  - Server is computationally unbounded, does not speak same language as User, but is “helpful”.
  - What kind of functions f?
    - E.g., uncomputable, PSPACE, NP, P?

Setup

User

Server

$f(x) = 0/1?$

$R \leftarrow \$\$\$$

q_1

Different from interactions in cryptography/security:

There, User does not trust Server, while here he does not understand her.

Computes $P(x,R,a_1,\ldots,a_k)$

Hopefully $P(x,\ldots) = f(x)$!

Intelligence & Cooperation?

- For User to have a non-trivial interaction, Server must be:
  - Intelligent: Capable of computing $f(x)$.
  - Cooperative: Must communicate this to User.
- Formally:
  - Server S is f-helpful if

    $\exists$ some (other) user U' s.t.

    $\forall$ x, starting states σ of the server

    $(U'(x) \leftrightarrow S(\sigma))$ outputs $f(x)$

Successful universal communication

- Universality: Universal User U should be able to talk to any (every) f-helpful server S to compute f.

- Formally:
  - U is f-universal, if
    $\forall$ f-helpful S, $\forall\ \sigma$, $\forall$ x
    $$(U(x) \leftrightarrow S(\sigma)) = f(x) \text{ (w.h.p.)}$$

- What happens if S is not helpful?
  - Paranoid view $\Rightarrow$ output "f(x)" or "?"
  - Benign view $\Rightarrow$ Don't care (everyone is helpful)

Main Theorems [Juba & S. '08]

- If f is PSPACE-complete, then there exists a f-universal user who runs in probabilistic polynomial time.
  - Extends to checkable ("compIP") problems
    - (NP ∩ co-NP, breaking cryptosystems)
    - S not helpful ⇒ output is safe
- Conversely, if there exists a f-universal user, then f is PSPACE-computable (in "compIP")
  - Scope of computation by communication is limited by misunderstanding (alone).

Proofs?

- Positive result:
  - f ∈ PSPACE ⇒ membership is verifiable.
  - User can make hypothesis about what the Server is saying, and use membership proof to be convinced answer is right, or hypothesis is wrong. Enumerate, till hypothesis is right.
- Negative result:
  - In the absence of proofs, sufficiently rich class of users allow arbitrary initial behavior, including erroneous ones.
  - (Only leads to finitely many errors …)

Implications

- Communication is not unboundedly helpful ☹
  - If it were, should have been able to solve every problem (not just (PSPACE) computable ones).
- But there is gain in communication:
  - Can solve more complex problems than on one's own, but not every such problem.
- Resolving misunderstanding? Learning Language?
  - Formally No! No such guarantee.
  - Functionally Yes! If not, how can user solve such hard problems?

Implications for Language Learning

- Well-explored theme in “linguistics”
  - Semantics learned by functional relevance.
  - But how does one have “common” grounding? Is this a purely a function of having common physical environment + needs?
  - Is there a purely intellectual basis for common grounding?
- Our answer: YES!

Towards Efficiency

- Learning of language is not efficient
  - User takes at least k steps to enumerate k possible servers (k possible languages).
  - Can this be made faster?

- Answers:
  - No! Not without assumptions on language …
  - Yes! If server and user are “broadminded”, and have “compatible beliefs” [JS ‘10]

Broadmindedness, Compatible beliefs:

- Beliefs of server S:
  - Expects users chosen from distribution X.
  - Allows “typical” user to reach goal in time T.
- Beliefs of user U:
  - Anticipates some distribution Y on users that the server is trying to serve.
- Compatibility: $K = (1 - |X - Y|_{TV})$
- Theorem[JS]: U can achieve goal in time poly(T/K).

Conclusions

- Basis of semantic communucation: Model "miscommunication"
  - Can be done by allowing users/servers to be variable (members of a set).
- Such settings seem commonplace, especially in "natural communication", but no prior attempts to model them theoretically (in the context of information transmission).
- Can also look at the "compression" problem.
  - Unveils phenomena reflective of natural communication [Juba,Kalai,Khanna,S. '10]

Thank You!