Communication as Coordination

Madhu Sudan
Microsoft, Cambridge, USA

-

Joint work with Jacob Leshno (MSR→Columbia)

Overarching question

- Can intelligent devices communicate intelligently?
  - Can printer explain to computer how to print on it?
- Devices communicate for some reason – some functionality
  - Can it check for errors in functionality?
  - It is interpreting the bits it receives correctly?
  - Are messages it is sending being interpreted correctly?
- Humans seem to manage:
  - How? Why?

Meaningful communication

- Meaning of bits
  - Bits sent/received are instructions
- Definition of Understanding?
  - Devices understands if device follows instructions?
    - Not appropriate if instructions are not aligned with devices functionality!
  - Need to know device's "incentives" and make sure they are compatible with instructions.
- Mixes communication with Game Theory!

(Repeated) Coordination Game

- Single instance
  - Alice and Bob have to simultaneously choose one of two actions, say {0,1}
  - If both pick same action, both win.
  - If they pick opposite actions, both lose.
- Main challenge: Don't know what the other person will choose when making our choice.
- Repeated version:
  - Play a sequence of games, using outcome of previous games to learn what the other player may do next.

Our setting

- Repeated coordination game with uncertainty:
  - Bob's perspective:
    - Knows his payoffs – 1 for coord.; 0 for not.
    - Does not know Alice's payoffs:
      - May vary with round
      - But for every action of Alice, payoff does not decrease if Bob coordinates (compatibility).
      - Knows a set S_A of strategies she may employ ("reasonable behaviors").
    - Can he learn to coordinate eventually?
      - After finitely many rounds of potential miscoordination, he should coordinate on every round forever.

Motivation

- Models natural communication: Communication aims to settle some coordination like problem.
  - People making choices; would like to be compatible with neighbors.
  - Other agents' motivations not completely clear to us.
  - Communication attempts to explain our intent; but we still have a choice on how to assign meaning to the bits.
  - (Mis)coordination (in last round) signals our current understanding may be (in)correct.

More formal definitions

- Strategy: (probabilistic) function from state and actions of last round to next state and action.
  - State can remember whatever we want about the past; in particular history of past plays.
  - S_A, S_B: strategy sets of Alice and Bob (public knowledge).
  - Alice plays some strategy in S_A, Bob in S_B.
  - Universality:
    - $\beta \in S_B$ is <u>universal</u> if $\forall\, \alpha \in S_A$, α and β reach eventual coordination in finite # of steps.

Coordinatable Strategies for Alice

- Set of possible Alice strategies can't be arbitrary:
  - Necessary condition (for universality):
    - $\alpha \in S_A$ is coordinatable if $\exists \beta \in S_B$ s.t. α and β reach eventual coordination.
  - Sufficient condition? Strong coordinatability:
    - $\alpha \in S_A$ is strongly coordinateable if $\exists \beta \in S_B$ s.t. $\forall$ states σ of Alice, (α, σ) reach eventual coordination with $(\beta, \emptyset)$.
- Main challenge remaining from previous works: Should our devices be coordinatable or universal?
  - Symmetry: Can we have $S_A = S_B =$ set of all reasonable strategies?

Theorems

- If S_A is enumerable and S_B $-$coordinateable, then $\exists S_B'$ that is enumerable such that
  - S_A is S'_B-coordinateable.
  - S_B' has a universal strategy for S_A

But $S_A \neq S_B'$ ☹

- If S_A is deterministic, $0 \leftrightarrow 1$-invariant, and S_A-coordinatable, then it does not have a universal strategy for itself.

Can't get $S_A \neq S_B'$ trivially ☹

- Corollary: No computable universal strategy against the set of computable strategies.

Theorems (contd.)

- (Main) If S_A is S_A-strongly-coordinatable and enumerable then $\exists\, S'_A \supseteq S_A$ which is also enumerable, $0 \leftrightarrow 1$-invariant, S'_A-strongly coordinateable and which has a universal strategy.
- So S'_A can be the model of “reasonable” strategies. We can use it to design our own universal protocols while allowing others to use other universal, but reasonable, protocols to learn us.
- Corollary: Example reasonable strategy set = probabilistic computable strategies.

Conclusions

- Game theory provides nice framework for casting questions about meaning of information.
- Using this framework, we can design communication strategies that learn to adapt and learn to serve functionality in a robust/changing environment.
- Need to make strategies “active-passive”
  - Should be capable of learning complex strategies and emulating them.
  - But should quickly revert to simple strategy that make it feasible for others to learn and emulate it.

Thank You