IMPLICIT COMMUNICATION IN A JOINT ACTION

Ross A. Knepper, Julia Proft, Christoforos I. Mavrogiannis, Claire Liang
Presented by Kiyun Kim
Definitions

- Implicit Communication: any expression of ideas through means other than explicit natural language statements
- Joint Activity: an activity performed by two or more agents in cooperation towards a common goal
  - A matter of the mental state of the involved agents
  - Participation may be asymmetric
Why Implicit Communication?

- People use implicit communication for a number of reasons—efficiency, social bonding, tact, and group cohesion—and will expect colleagues (human or robotic) to be able to comprehend it.

- If a robot is not cognizant of how humans interpret actions through implicit communication, it will be perceived as sending random implicit signals.

- Human interaction and perception is governed by a common set of assumed functional and social capabilities, but most humans are uninformed about a robot’s capabilities.

- A robot that fails to work as expected is unlikely to remain in use, even if the failure comes from human expectations, rather than the robot’s functional capabilities.
Communication in Joint Action

- Any action which implicitly communicates information is perceived by the observer with some measure of surprisal, an encoding of how probable the observer believes the action to be given the context.
  - Higher surprisal indicates that the action is more improbable within the context.
  - Greater surprisal usually corresponds to a more meaningful or strongly-conveyed action.
Communication in Joint Action

- $\mathcal{A}$: the set of all possible actions
- $\mathcal{M}$: the set of all possible facts about the world
- An agent performs a series of actions $a_1, a_2, \ldots, a_n \in \mathcal{A}$
- $A_i \subseteq \mathcal{A}$ is the set of possible ways of accomplishing the functional goal of an action (i.e. a subgoal of the joint activity)
- An agent $Q$ performs actions in a context $M_Q \subseteq \mathcal{M}$, the set of facts capturing $Q$’s individual knowledge.
  - Includes common-sense knowledge, state history of all agents, the observable scene, and properties of objects within the scene.
  - Each $m \in M_Q$ can have associated confidence to allow for changes
  - $M_Q^{\text{priv}}$ is agent $Q$’s private knowledge
- $M_{\text{pub}}$ is the set of all public knowledge
  - $m \in M_{\text{pub}}$ unless all agents are aware that $m$ is knowledge shared by all.
- The distribution $P(a | M)$ describing the likelihood that a specific agent may next perform action $a$ in the specific context $M$. 
Communication in Joint Action

- $\mathcal{A}$: the set of all possible actions
- $\mathcal{M}$: the set of all possible facts about the world
- The goal for agent Q and observer R is for Q to select an action $\hat{a} \in A_f$ that is perceived by R to be improbable within a given context and communicates $\hat{m} \in M^Q_{\text{priv}}$.
- R then seeks an explanation for $a'$ in the form of a previously unknown fact $m'$, which makes $a'$ less surprising.
- This goal is not always possible.
- In order for R to correctly interpret Q’s meaning, $\hat{a}$ and $\hat{m}$ must meet four criteria for implicit communication.
Implicit Communication Criteria

1. \( \exists \hat{a}, a' \in A_f: \hat{a} \neq a' \)
   - There must be at least two feasible, distinct actions that accomplish the functional goal.

2. \( P(\hat{a}|M_{pub}) < P(a'|M_{pub}) - \varepsilon \)
   - There should be a more obvious choice of action, \( a' \), in order to prompt the observer to search for an explanation of why the actor chose action \( \hat{a} \).

3. \( P(\hat{a}|M_{pub}) < P(\hat{a}|\hat{m}, M_{pub}) - \varepsilon \)
   - The fact \( \hat{m} \) should be easy for the observer to identify as an explanation of \( \hat{a} \).

4. \( \forall m \in M \setminus M_{pub} \cup \{\hat{m}\}: P(m|\hat{a}, M_{pub}) < P(\hat{m}|\hat{a}, M_{pub}) - \varepsilon \)
   - No other inferred meaning \( m \) is more likely than the intended explanation \( \hat{m} \).
Understanding

- Upon seeing action $\hat{a}$, R infers $\hat{m} \leftarrow \arg\max_{m \in M} P(m | \hat{a}, M_{\text{pub}})$

- Using Bayes’ Rule:

\[
\hat{m} \leftarrow \arg\max_{m \in M} \frac{P(\hat{a} | m, M_{\text{pub}})P(m | M_{\text{pub}})}{P(\hat{a} | M_{\text{pub}})} = \arg\max_{m \in M} P(\hat{a} | m, M_{\text{pub}})P(m | M_{\text{pub}}).
\]
Generating

- Identical to the process for understanding, but we search over the set of actions instead of the set of actions.

\[
\hat{a} \leftarrow \arg\max_{a \in A_f} P(\hat{m} | a, M_{pub})
\]

\[
\hat{a} \leftarrow \arg\max_{a \in A_f} \frac{P(a | \hat{m}, M_{pub}) P(\hat{m} | M_{pub})}{P(a | M_{pub})}
\]

\[
= \arg\max_{a \in A_f} \frac{P(a | \hat{m}, M_{pub})}{P(a | M_{pub})}
\]
Achieving Joint Goals

- Trust – confidence in another agent’s future actions
  - Limited by common understandings amongst agents, even when all are behaving rationally
  - Also limited by discrepant beliefs about facts
- Consensus – all agents preferring the same joint strategy S to accomplish a goal
  - A group of competent agents who are able to implicitly communicate can achieve consensus
- Receptivity – how receptive each agent is to updating their beliefs about the joint strategy
  - Highest when no consensus for a joint strategy has been made
  - Once consensus has been made, agents will operate under civil inattention
Case Study: Natural Language

- Conversational Implicature
  - Should be cancelable—i.e. there exists some phrase that, when appended to the end of the sentence, cancels the meaning of implication by explaining it
  - Gricean Maxims: Quantity, Quality, Relation, and Manner
  - Deliberately breaking a maxim is an improbable action and shows the use of implicature

- Inverse Semantics
  - Speech entailment
  - Generalized Grounding Graph ($G^3$)
Case Study: Motion

- Legibility – weighted sum of the probabilities that observers assign to the actor’s intended goal throughout the whole trajectory
  - Higher legibility means a higher bias towards predicting the agent’s actual goal

- Dynamic Legibility
  - Maxims for motion: Efficiency, Motion, and Manner
  - Enforcing these maxims leads towards consensus
  - Once consensus is reached, civil inattention.
  - Pick actions that decrease uncertainty in order to reach consensus
Implementation

- Existing Implementations: AI search-pruning techniques to restrict the action space $A_f$
  - $A_f$ is often hard-coded, meaning that it might not match human expectation
- Complete model of $M_{pub}$ is often unnecessary and infeasible – it is frequently less costly to repair a misunderstanding.
- $P(a|M)$ is best modeled through machine learning – the ways a particular context impacts the probabilities of various actions are complex