The Communicative Multiagent Team Decision Problem

Analyzing Teamwork Theories and Models

Presented by: The Cool Team
Aaron Brodersen, Brian Rezac, and David Montz

Citations

Pynadath, D. and M. Tambe (2002). The Communicative Multiagent Team Decision Problem: Analyzing Teamwork Theories and Models, *Journal of artificial Intelligence,* **1**(6):389-423.

Tambe, M. (1997). Towards flexible teamwork. *Journal of artificial Intelligence Research, 7,* 83-124.

Jennings, N. (1995) Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence, 75,* 195-240.

Overview

- The COM-MTDP Model
- Theorems and Complexity Analysis
- Framework Demonstration
- Empirical Results

Background

- Research in multiagent teamwork doesn’t address the optimality or complexity of the teamwork problem
- The Communicative Multiagent Team Decision Problem (COM-MTDP) is put forth as a tool to evaluate such tradeoffs.

COM-MTDP

- Inspired by work in economic team theory
- Meant to address the shortcomings of belief-desire-intention (BDI) frameworks of cooperation, among others.
- Quantitatively evaluates the optimality of coordination behavior
- Characterizes the computation complexity of the aspects of team decisions

The COM-MTDP Model

Ingredients:

- World state matrix **S**
- Domain level actions (agent decisions) $\mathbf{A}_{\alpha}$
- A probability distribution **P** for state transitions caused by actions
- Agent observations **Ω** (includes all past observations)

The COM-MTDP Model

Ingredients Continued:

- Uncertainty in observations (stochastic) **O**
- Agent communication Σ
- Agent beliefs **B** resulting from observation and communication
- A reward function **R** of state, actions, and communication

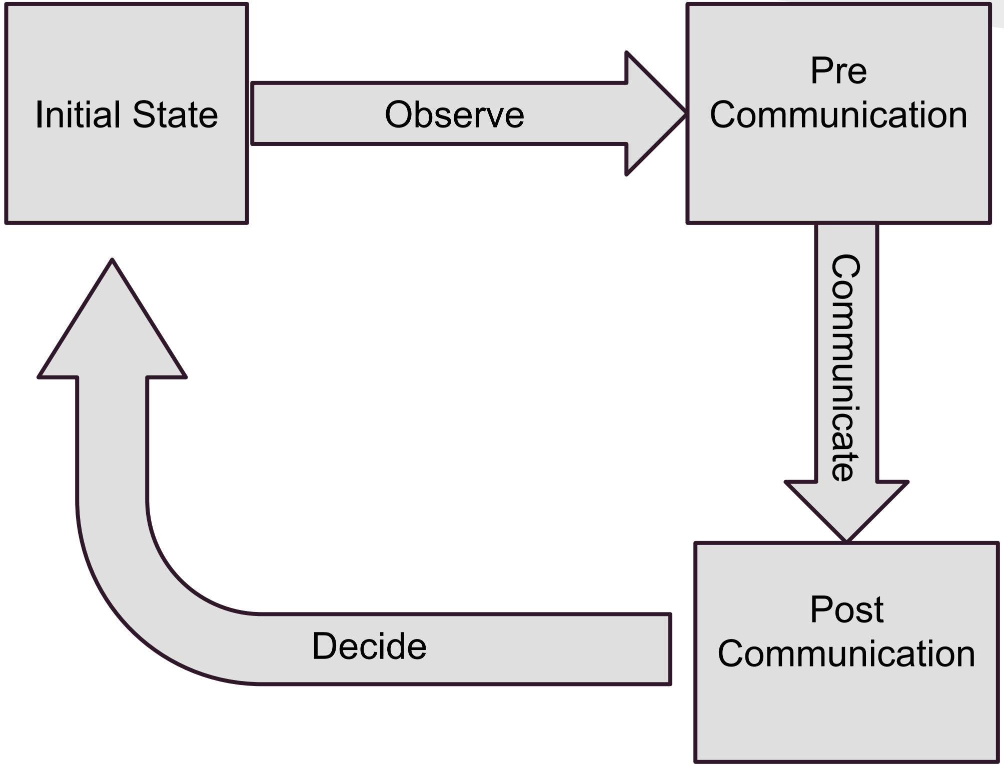
Observability Assumptions

- Stochastic
- Perfect Recall
- Observability levels
  - **None** All agents are blind
  - **Collective-Partial** The team can see part of the world
  - **Collective** The team can see the entire world
  - **Individual** Each agent can see the entire world

Communication Assumptions

- Instantaneous
- Non-intermittent (no noise or latency)
- Perfect recall
- Non-negative cost
- Communication levels
    - Normal
    - Free (No Cost)
    - None (Radio Silence)

Agent States



Initial State

Observe

Pre Communication

Communicate

Post Communication

Decide

Paper mainly focuses its analysis on agent communication policies

Theorems

- Lots of theorems and proofs
- Boil down to these facts
  - Under free communication, the action selection policy will optimize the utility
  - Determining the existence of optimal action and communication policies is a decision problem
  - The complexity of this problem depends on observability and communication cost

Policy Existence Complexity

Communication	Observability			
	Individual	Collective	Collective-Partial	None
None	P-complete	NEXP-complete	NEXP-complete	NP-complete
General	P-complete	NEXP-complete	NEXP-complete	NP-complete
Free	P-complete	P-complete	PSPACE-cmplt	NP-complete

Each complexity class is a subset of those to the right			
P	NP	PSPACE	NEXP
polynomial time	non-deterministic polynomial time	polynomial memory	$O(2^{polynomial})$

Demonstration and Analysis

- Prove out the COM-MTDP framework
    - Analysing several algorithms
    - Calculate empirical results of algorithms
    - Comparing the results

Demonstration and Analysis

- Joint Intention Theory
- Three Implementations for comparison
    - Jennings implementation
    - STEAM
    - Local Optimum
- Little justification for this choice

Jennings Implementation

- Doesn’t factor in communication cost when sending a message
- Always sends a message when the state of the team’s goal changes

STEAM

- Send a message when this simple inequality is true

$$\tau \cdot C_{mt} > C_C$$

- Uses simple parameters to specify observability and communication cost
- Compare cost of communicating (C_C) against the cost of miscoordination (C_{mt}) adjusted for observability (τ)

Local Optimum

- Used to demonstrate best possible choice an agent could make
- Hard to calculate in partially observable, non-free communication environments
- Not a usable communication policy in most problems

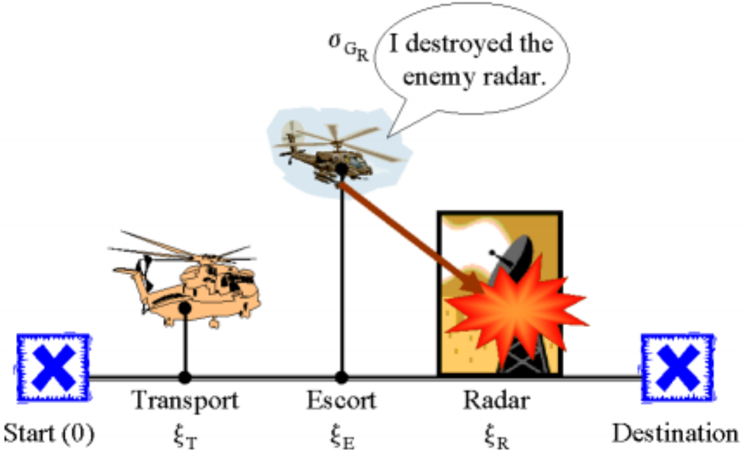
Local Optimum Complexity

$|S|$ = number of world states $|\Omega_a|$ = number of agent observations T = time

	Observability			
Communication	Individual	Collective	Collective-Partial	None
None	$\Omega(1)$	$\Omega(1)$	$\Omega(1)$	$\Omega(1)$
General	$\Omega(1)$	$O((\lvert S\rvert \cdot \lvert\Omega_a\rvert)^T)$	$O((\lvert S\rvert \cdot \lvert\Omega_a\rvert)^T)$	$\Omega(1)$
Free	$\Omega(1)$	$\Omega(1)$	$\Omega(1)$	$\Omega(1)$

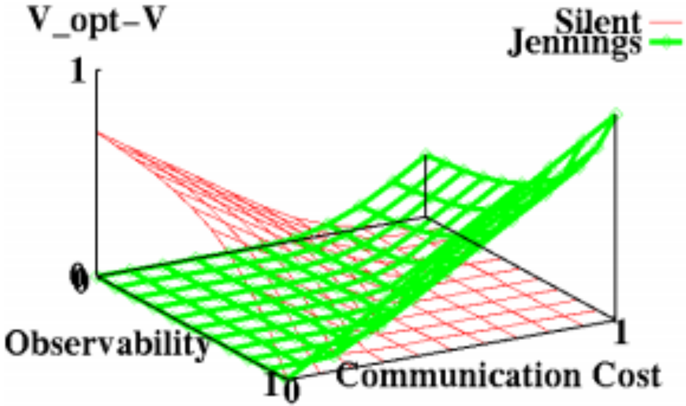
Experiment Setup

- Escort: Can evade and destroy radar
- Transport: Cannot evade or destroy radar, must fly slower until the radar is known to be destroyed
- Radar: Placed Randomly
- Goal: Travel from start to destination in minimal time



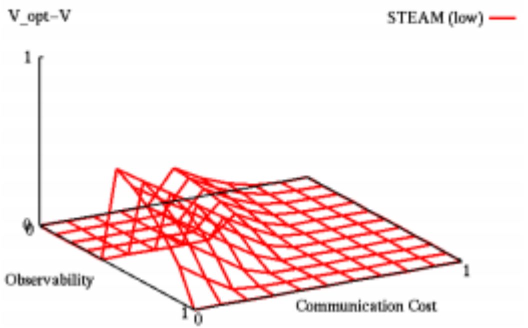
Jenning’s Policy

- Sends the message in all circumstances
- Jenning’s under no communication cost is optimal
- However as the cost of communication rises it becomes sub-optimal
- In high communication costs silence becomes more optimal than Jenning’s



STEAM Policy with Low C_{mt}

- STEAM underperforms at no communication cost and medium observability and at no observability and medium communication cost
- STEAM provides optimal performance in all edge cases
- STEAM handles communication costs better than Jennings



STEAM Policy with Medium C_{mt}

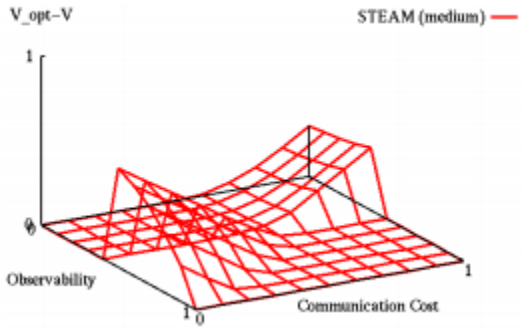
- STEAM with medium communication cost under performs at low observability and high communication cost and medium observability and low communication

  costs
- Steam performs optimally under high

  observability and under low observability

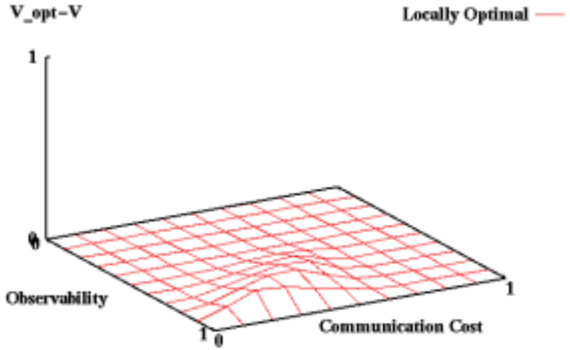
  with low communication costs
- STEAM handles communication

  costs better than Jennings



Local Optimality Policy

- The local optimality policy functions near optimally under all conditions except high observability with low communication costs
    - But even its least favorable point is more optimal than most of STEAM or Jenning’s



Local Optimality Benefits

- While Local Optimality Policy is not guaranteed optimality it is far less resource intensive than the Global Optimality
    - 5 seconds as opposed to 150 minutes

- On average Local Optimality only varied 1.1% from Global Optimality, and at most it varied 12%

Conclusion

- Under non-zero communication costs Jenning's Policy can be inferior to a silent policy
- STEAM under low or medium cost have large segments of non-optimality near medium observability
- The local optimality policy produces very optimal results over the entire domain of observability and communication costs

Critiques

- The given test, while a good way to test the COMP-MTDP model is overly simplistic

- Local Optimality might only do so well because escort can perfectly observe the transition of the transport's state if it waits one time step
    - Not all problems have such easily observable states

Future Works

- Adjusting COM-MTDPs to handle unknown outcome probabilities or rewards

- Adding team formation as a complement to the COMP-MTDPs communication

- Relaxing the requirement that all agents in a COMP-MTDPs are selfless and share a utility

Questions?