Multiagent Planning under Uncertainty with Stochastic Communication Delays

Matthijs Spaan[§] Frans Oliehoek[*] Nikos Vlassis[#]

[§]Institute for Systems and Robotics, IST, Lisbon, Portugal

[*]University of Amsterdam, The Netherlands

[#]Technical University of Crete, Chania, Greece

Lisbon Workshop on New Challenges for Cooperative Robotics

October 24, 2008

Introduction

Focus:

- Cooperative multiagent planning under uncertainty.
- Formalized as a decentralized partially observable Markov decision process (Dec-POMDP).
- Very hard to solve.

Communication:

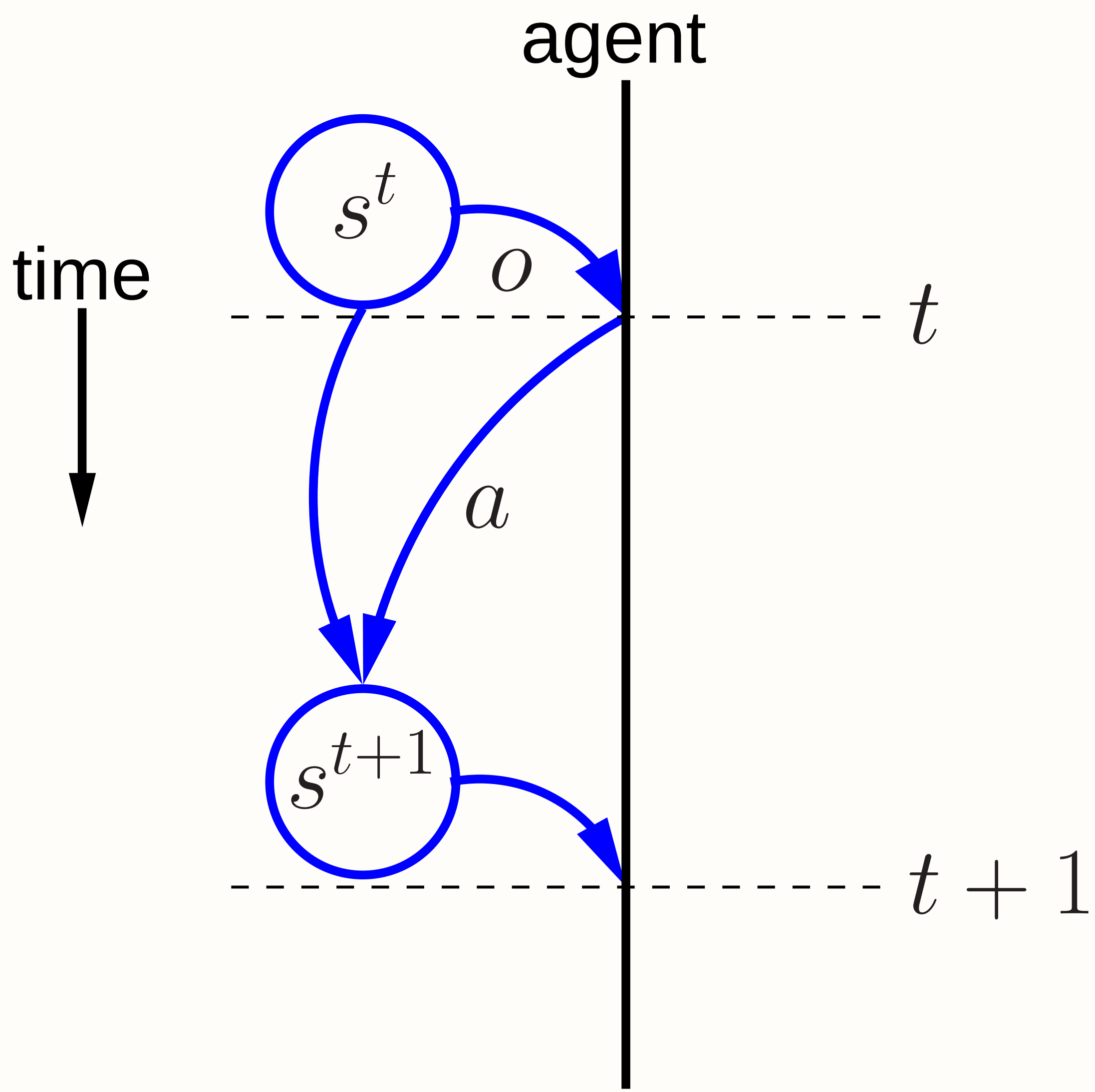
- When available can be leveraged.
- Sharing observations $\rightarrow$ synchronized knowledge.
- But requires instantaneous, perfect communication.

Contributions:

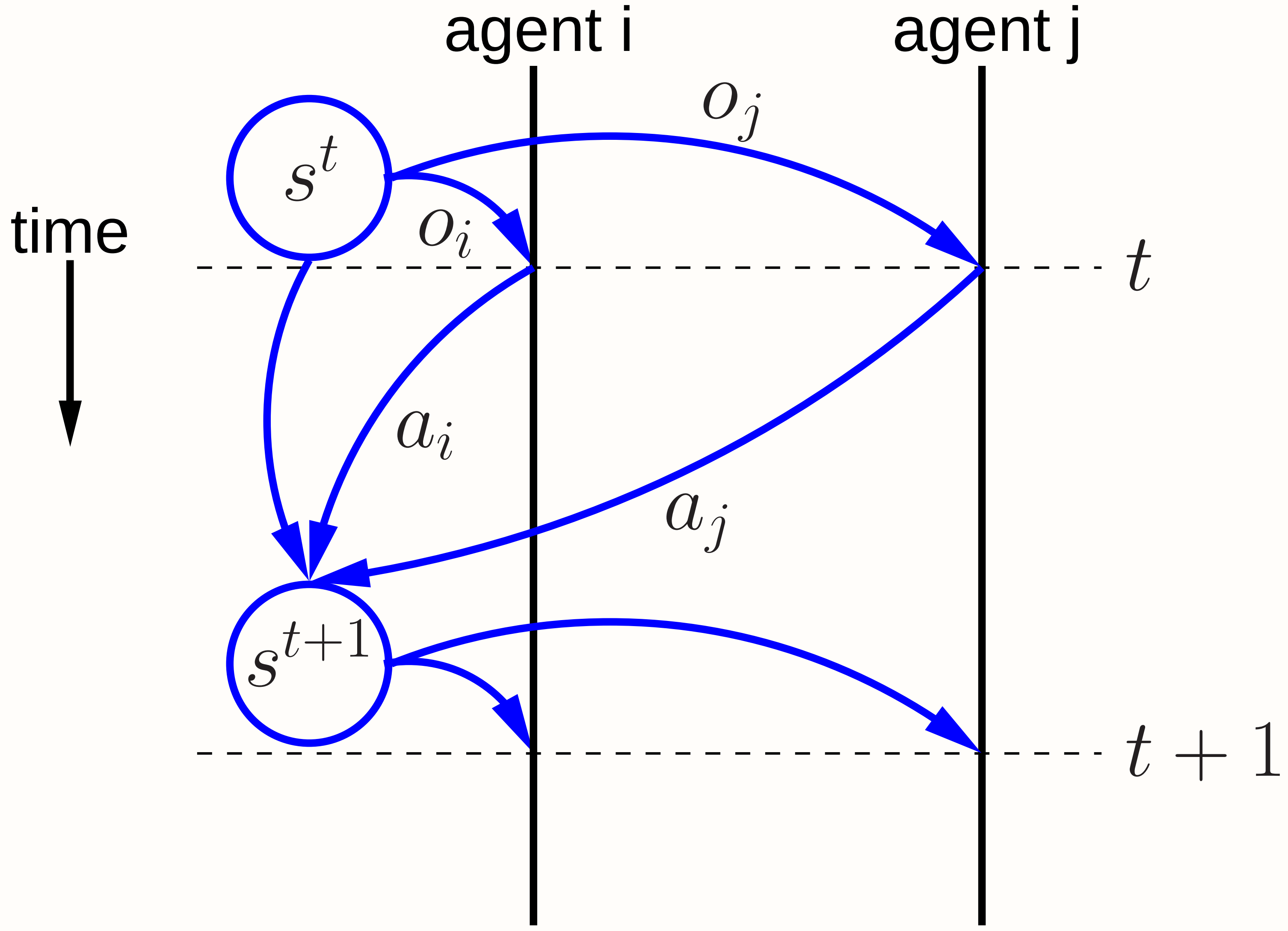
- We consider stochastic delays: towards more realistic communication models for Dec-POMDPs.

Outline

- POMDPs vs. Dec-POMDPs.
- Review of communication models:
  - Instantaneous, perfect communication.
  - Delayed communication (fixed delay).
- Stochastic communication delays:
  - Model.
  - Compactly representing the value function.
  - Value iteration method.
  - Experimental results.
- Conclusions.


agent
time
s^t
o
t
a
s^{t+1}
$t+1$

Dec-POMDP (without communication)



POMDPs vs. Dec-POMDPs

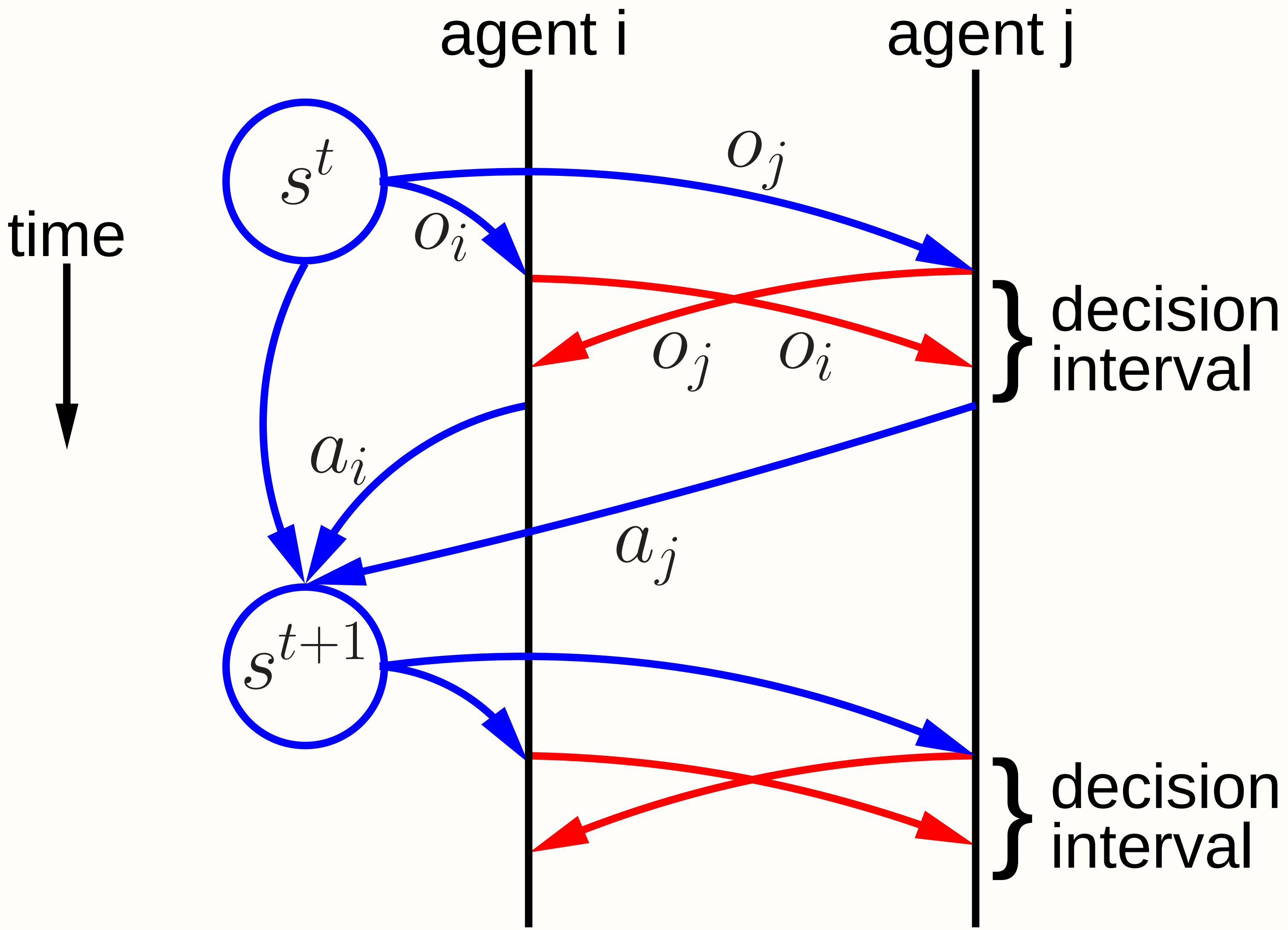
Frameworks for acting optimally given:

- limited sensing,
- stochastic environments.

Dec-POMDPs:

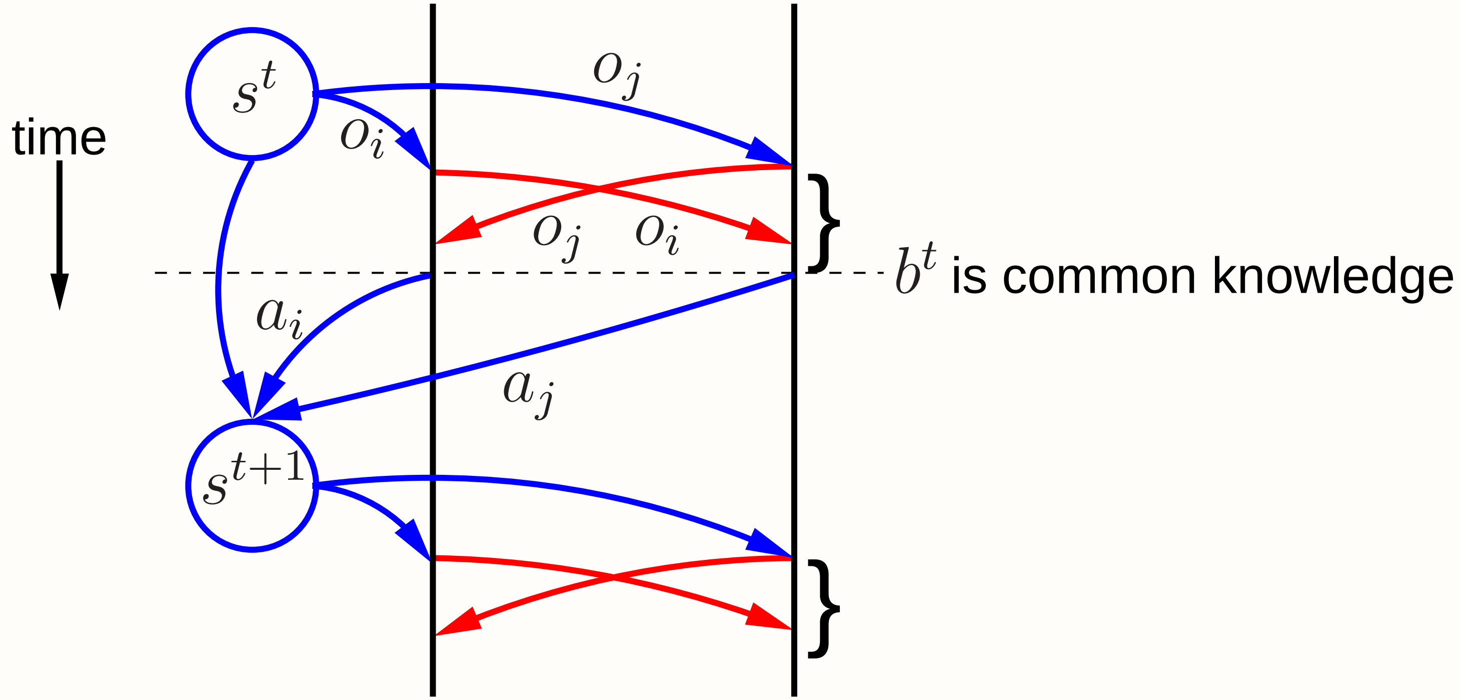
- Decentralized execution.
- Usually centralized planning.
- No common state estimate, no joint belief.
- Optimal policies based on observation history mapping.

Instantaneous communication



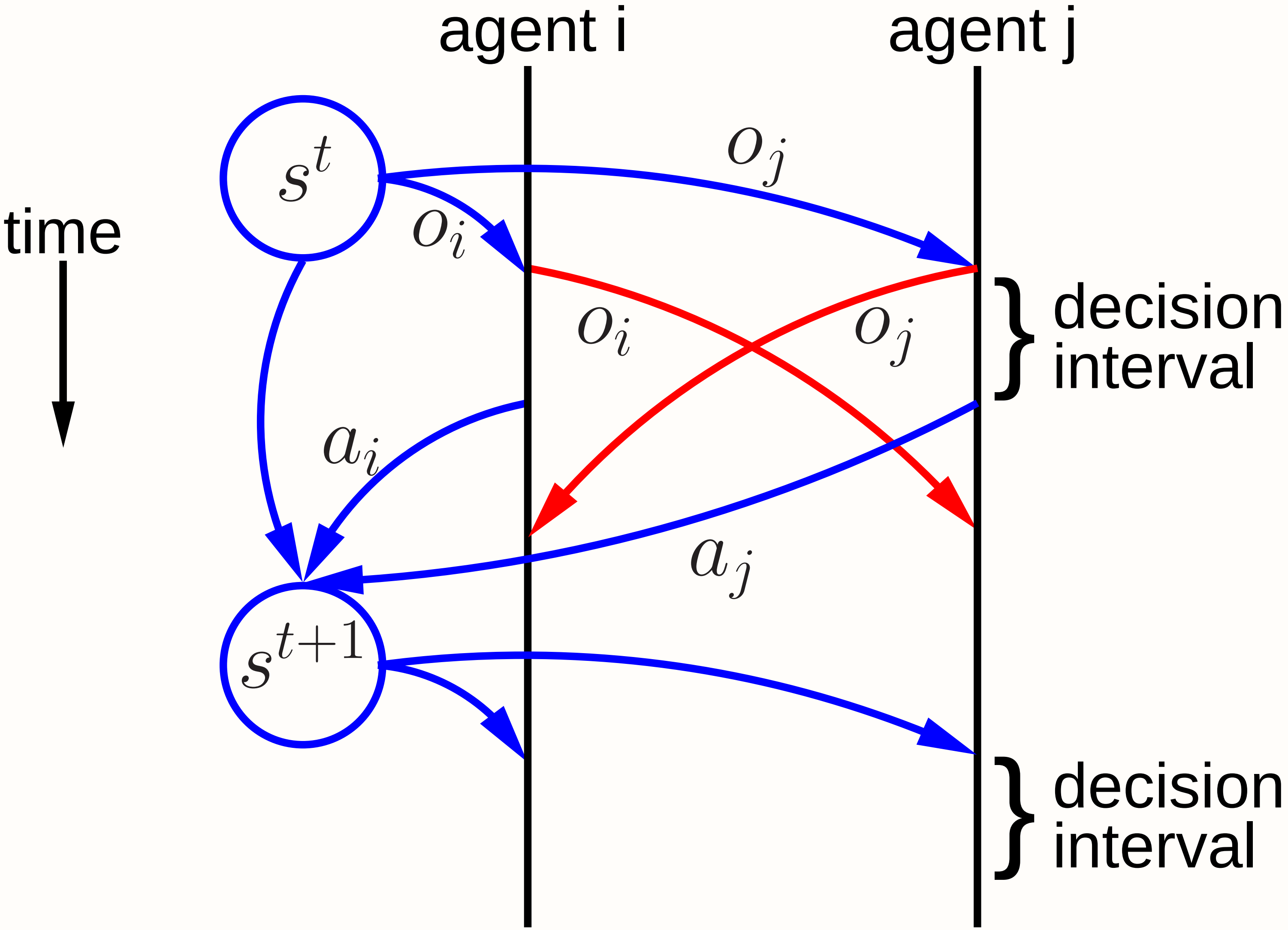
Instantaneous communication

- Instantaneous, perfect communication reduces a Dec-POMDP to a POMDP (Pynadath and Tambe, 2002).
- Observation sharing $\rightarrow$ each agent knows belief b^t at time t.
- Piecewise-linear and convex (PWLC) value function.
- But this synchronization step takes time.



INSTITUTO SUPERIOR TÉCNICO

Delayed communication



PÓLO DO I.S.T

Fixed delay communication

- Each agent knows the last commonly known b^{t-k}, $k > 0$.
- If $k = 1$, fixed delay communication can be modeled as a Bayesian game (Oliehoek et al., 2007).
- If $k = 1$, the optimal value function is PWLC, and equal to the Q_{BG} value function (Oliehoek et al., 2008).
- Otherwise the value function is not separable (Varaiya and Walrand, 1978), it is not a function over the joint belief space.
- But what if the delay can vary?

INSTITUTO SUPERIOR TÉCNICO

Stochastic communication delays

- We formalize the probability that synchronization succeeds within a particular stage (details in Spaan et al. (2008)):

$p^{0\mathrm{TD}}(s)$	synchronization is instantaneous
$p^{1\mathrm{TD}}(s)$	synchronization succeeds within 1 stage
$p^{2\mathrm{TD}}(s)$	synchronization succeeds within 2 stages
$\vdots$	$\vdots$

- Can depend on state, and $\sum_i p^{i\mathrm{TD}}(s) = 1$.
- E.g., robots close $\longrightarrow$ more reliable communication.
- Unifying framework:
  - $p^{0\mathrm{TD}}(s) = 1, \forall s$: instantaneous communication.
  - $p^{1\mathrm{TD}}(s) = 1, \forall s$: one-step delay communication.

PÓLO DO I.S.T

Stochastic communication delays

- Optimal value function:

$$Q^*_{\mathrm{SD}} = R + p^{0\mathrm{TD}} F_{0\mathrm{TD}} + p^{1\mathrm{TD}} F_{1\mathrm{TD}} + p^{2\mathrm{TD}} F_{2\mathrm{TD}} + \ldots,$$

  where F_{iTD} is the expected future reward given a delay of i stages.

- We assume during planning that delay is at most 1 step:

$$p^D = p^{1\mathrm{TD}} + p^{2\mathrm{TD}} + \cdots = 1 - p^{0\mathrm{TD}},$$

  and define an approximate value function as

$$\widetilde{Q}^*_{\mathrm{SD}} = R + p^{0\mathrm{TD}} F_{0\mathrm{TD}} + p^D F_{1\mathrm{TD}}.$$

- We prove that $\widetilde{Q}^*_{\mathrm{SD}}$ is PWLC over the joint belief space.

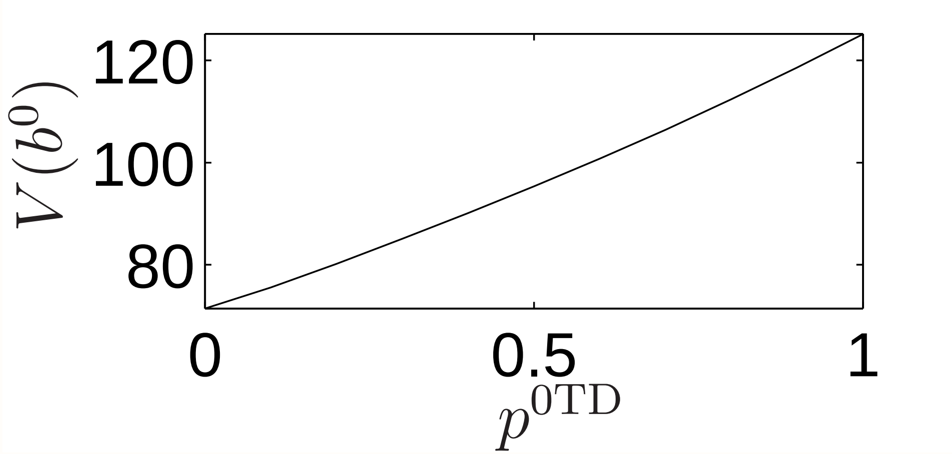
Algorithm

- We show how to perform finite-horizon value iteration for the stochastically delayed communication setting.
- Point-based approximate POMDP techniques transfer directly to our $\widetilde{Q}^*_{\mathrm{SD}}$ value function.
- For delays of more than 1 one time step, we propose an online algorithm similar to Dec-COMM (Roth et al., 2005).

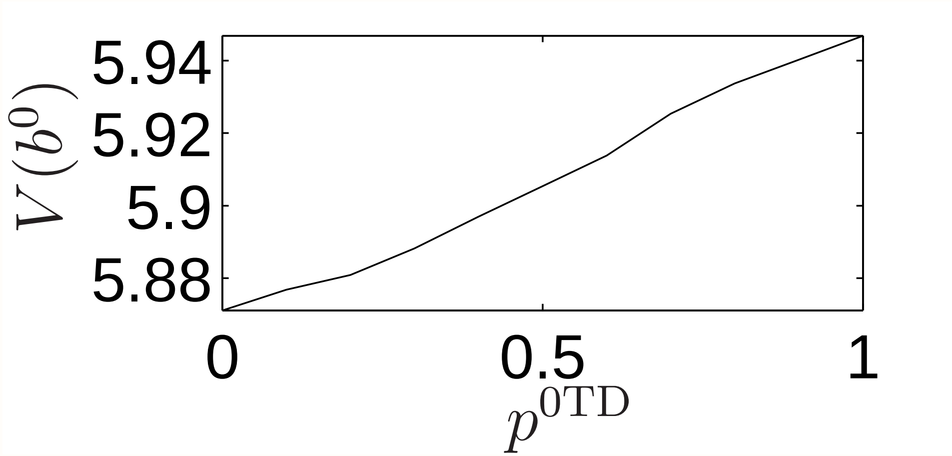
Experimental results

Uniform delay between 0 and 1:

Dec-Tiger

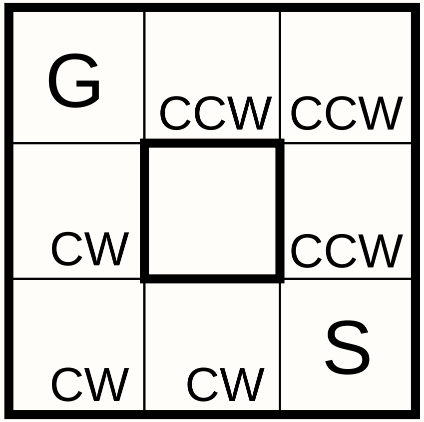


GridSmall



Experimental results

Meet in corner

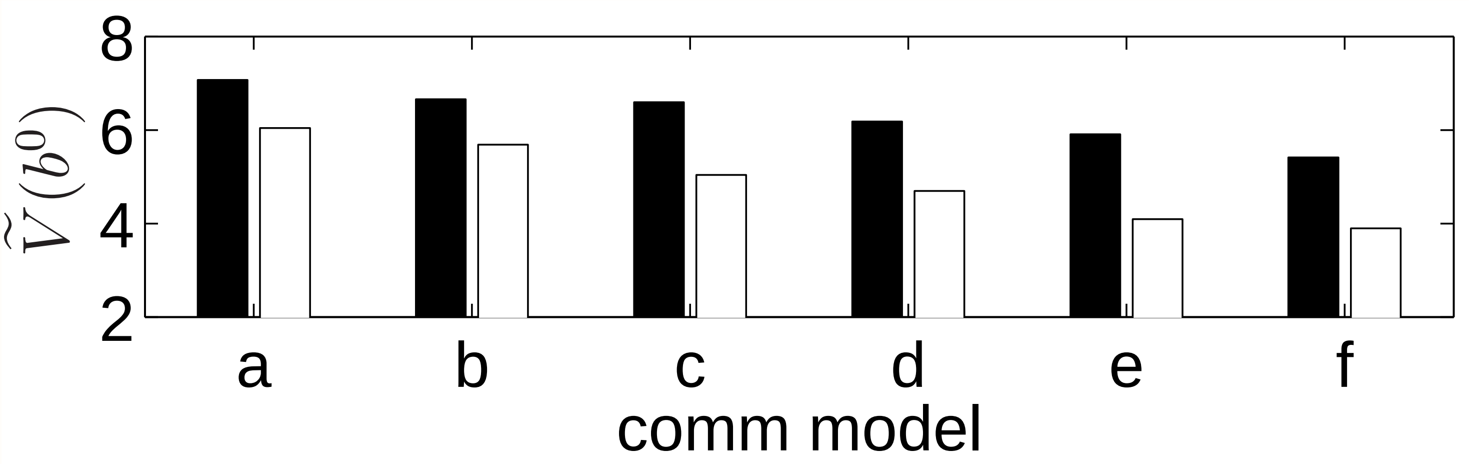


- The policies consider potential future communication capabilities.
- The CW route is cheaper, and with uniform $p^{0\mathrm{TD}}$ it is chosen.
- We set $p^{0\mathrm{TD}}(s) = 0, \forall s \in \mathrm{CW}$ and to 1 everywhere else.
- Now the agents choose the CCW route, leading to higher payoff.

Experimental results

Arbitrary delays:

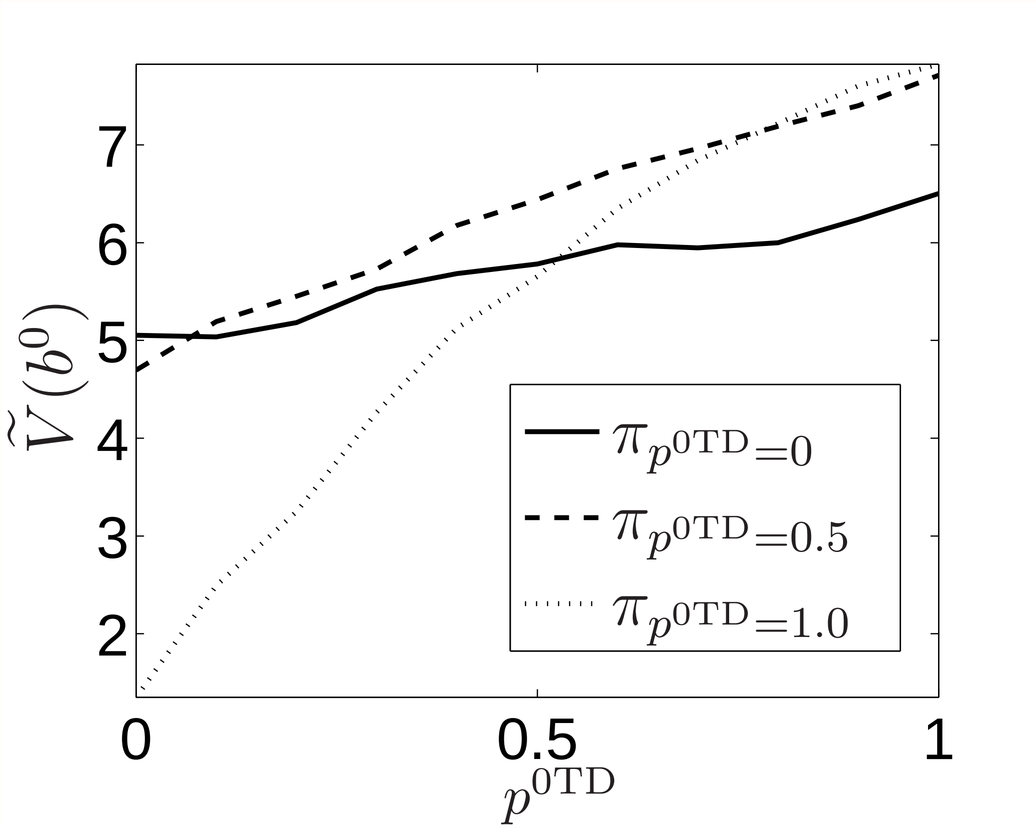
	a	b	c	d	e	f
$p^{0\text{TD}}$	0.8	0.8	0.6	0.6	0.6	0.6
$p^{1\text{TD}}$	0.2	0.1	0.4	0.3	0.2	0.2
$p^{2\text{TD}}$		0.1		0.1	0.2	0.1
$p^{3\text{TD}}$						0.1



Experimental results

Policies computed for a specific value of $p^{0\text{TD}}$, but evaluated under different values of $p^{0\text{TD}}$.

Meet in corner (large)



Conclusions

Conclusions:

- We discussed different communication assumptions made in the Dec-POMDP literature.
- We introduced a model for stochastically delayed communication:
  - More realistically models unreliable communication.
  - We showed how we can approximate its optimal value function, and how to compute it efficiently.
- Encouraging experimental results.

Future work:

- Not only delayed, but also noisy communication.
- Alternative methods for longer delays.
- Infinite-horizon value iteration.

INSTITUTO SUPERIOR TÉCNICO

References

F. A. Oliehoek, M. T. J. Spaan, and N. Vlassis. Dec-POMDPs with delayed communication. In *Multi-agent Sequential Decision Making in Uncertain Domains*, 2007. Workshop at AAMAS07.

F. A. Oliehoek, M. T. J. Spaan, and N. Vlassis. Optimal and approximate Q-value functions for decentralized POMDPs. *Journal of Artificial Intelligence Research*, 32:289–353, 2008.

D. V. Pynadath and M. Tambe. The communicative multiagent team decision problem: Analyzing teamwork theories and models. *Journal of Artificial Intelligence Research*, 16:389–423, 2002.

M. Roth, R. Simmons, and M. Veloso. Reasoning about joint beliefs for execution-time communication decisions. In *Proc. of Int. Joint Conf. on Autonomous Agents and Multi Agent Systems*, 2005.

M. T. J. Spaan, F. A. Oliehoek, and N. Vlassis. Multiagent planning under uncertainty with stochastic communication delays. In *Proc. of Int. Conf. on Automated Planning and Scheduling*, pages 338–345, 2008.

P. Varaiya and J. Walrand. On delayed sharing patterns. *IEEE Transactions on Automatic Control*, 23(3):443–445, June 1978.

PÓLO DO I.S.T