Decision-Making Under Uncertainty in Multi-Agent and Multi-Robot Systems: Planning and Learning

Chris Amato
Northeastern University

Northeastern University
College of Computer and Information Science

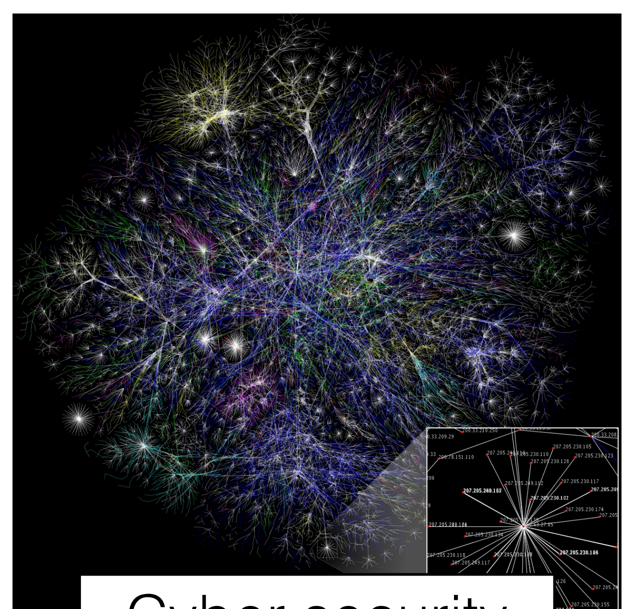
Multi-agent systems are (going to be) everywhere

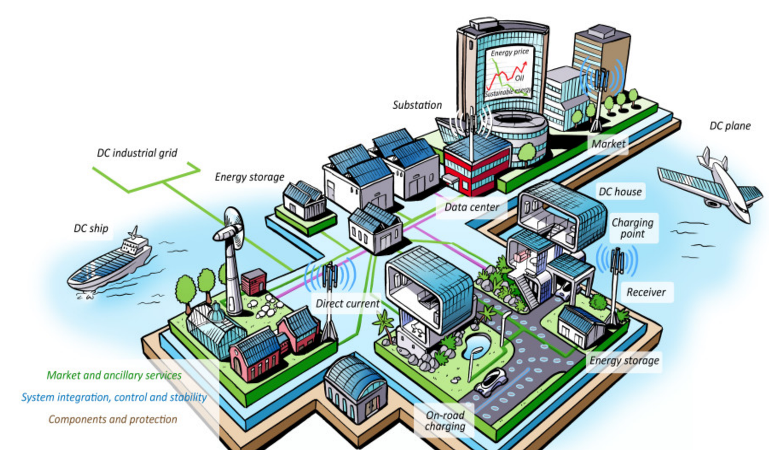
Drone delivery


UAV surveillance


Autonomous cars

Search and rescue


Cyber security



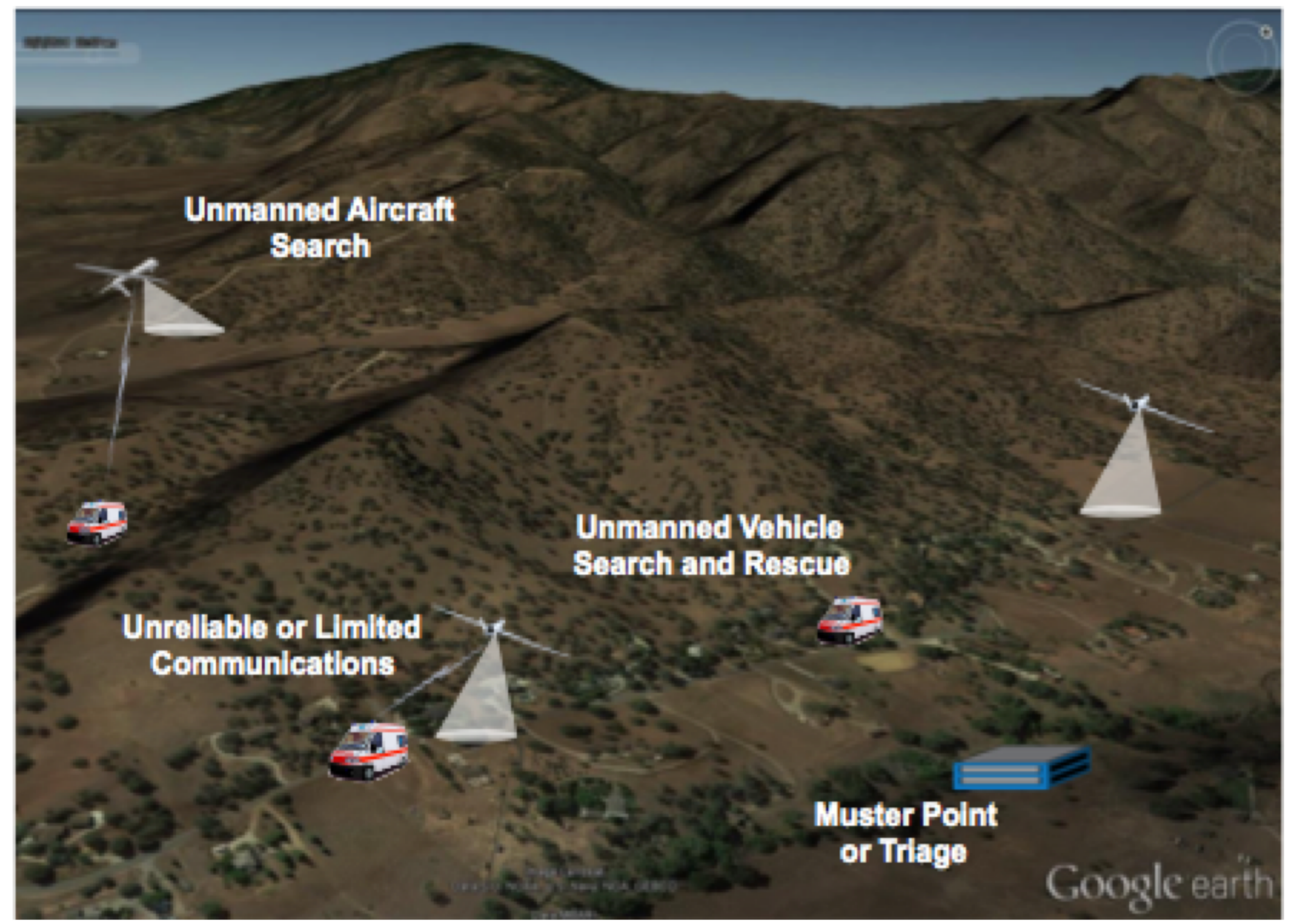
Smart energy grids

Uncertainties

- These real-world problems have several forms of uncertainty:
    - Outcome uncertainty
    - Sensor uncertainty
    - Communication uncertainty

Uncertainties: Search and Rescue

- A team of ground and aerial robots searching for people after a disaster:



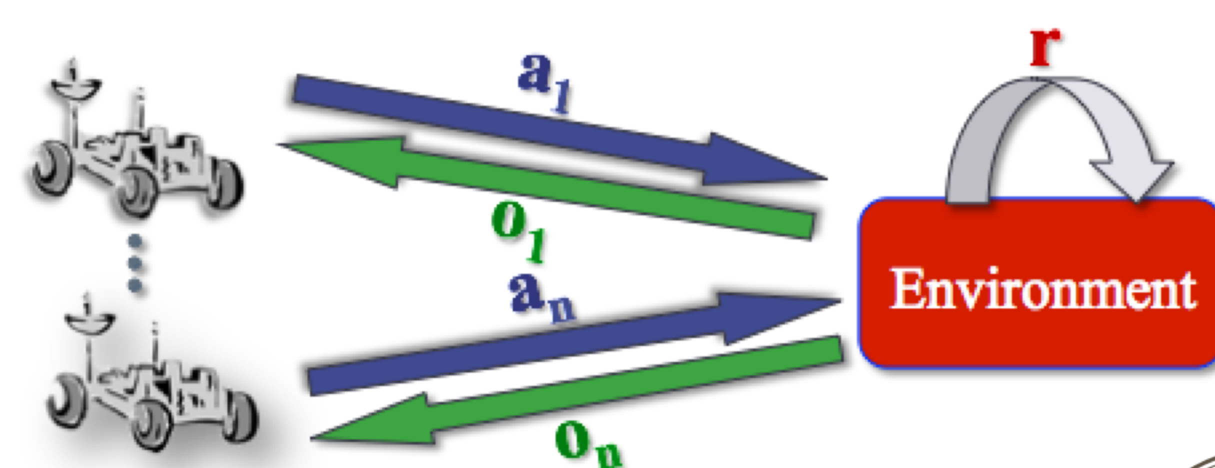
  - Outcome uncertainty: movement of robots and people is uncertain
  - Sensor uncertainty: location of people and obstacles is uncertain
  - Communication uncertainty: communication range is limited, so location and choices of other robots is uncertain

Uncertainties

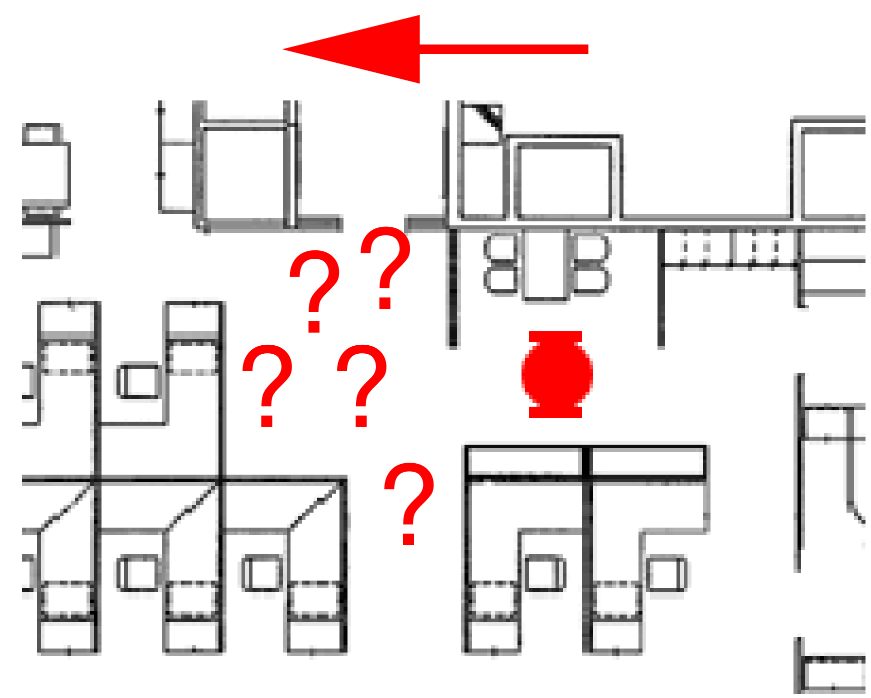
- Many other real-world problems have outcome, sensor and communication uncertainty

Multiple cooperating agents

- Decentralized partially observable Markov decision process (Dec-POMDP) Bernstein et al., 02
    - Extension of the single agent MDP and POMDP models
    - Multiagent sequential decision-making under uncertainty
- At each stage, each agent takes an action and receives:
    - A local observation
    - A joint immediate reward



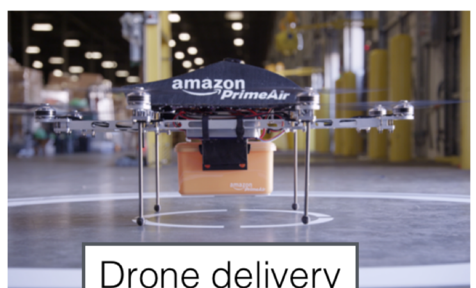
Dec-POMDP



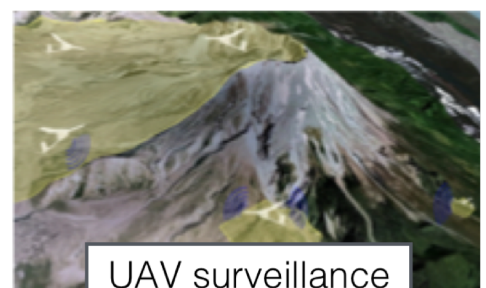
- Decision-theoretic model of multi-agent systems
- We need to model:
  - Each agent's actions, A_i, for each agent i
  - The environment states, S
  - The environment dynamics, $\Pr(s'|s, a)$
  - Each agent's sensor and communication observations, Ω_i
  - The observation function, $\Pr(o|s', a)$
  - The reward function, $R(s, a)$
- A solution seeks to maximize the expected sum of rewards from policies that only consider *local* observations

Dec-POMDPs are general

- Any real-world problem with outcome, sensor and communication uncertainty

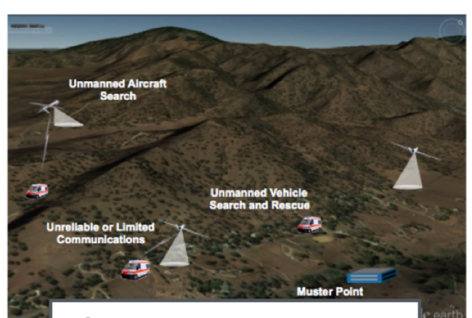


Drone delivery

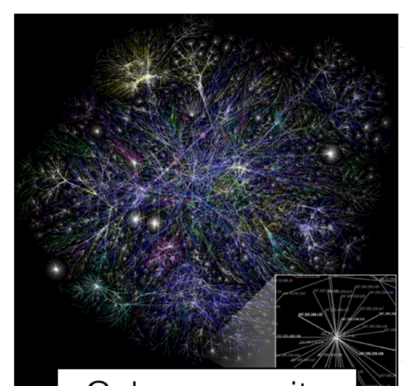


UAV surveillance

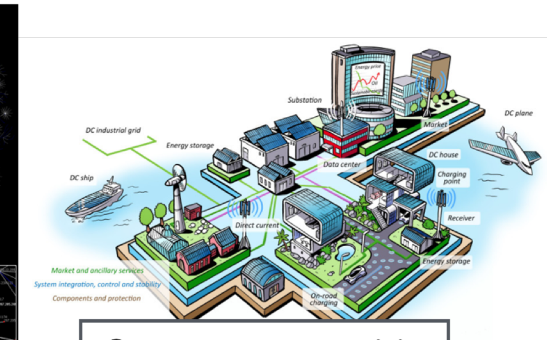
Autonomous cars



Search and rescue



Cyber security



Smart energy grids

- If we can solve the Dec-POMDP optimally, we get an optimal solution to our problem

Now what?

- Any cooperative multi-agent problem is a Dec-POMDP
- But, modeling and solving is hard
- Solutions: approximate the model or approximate the solution

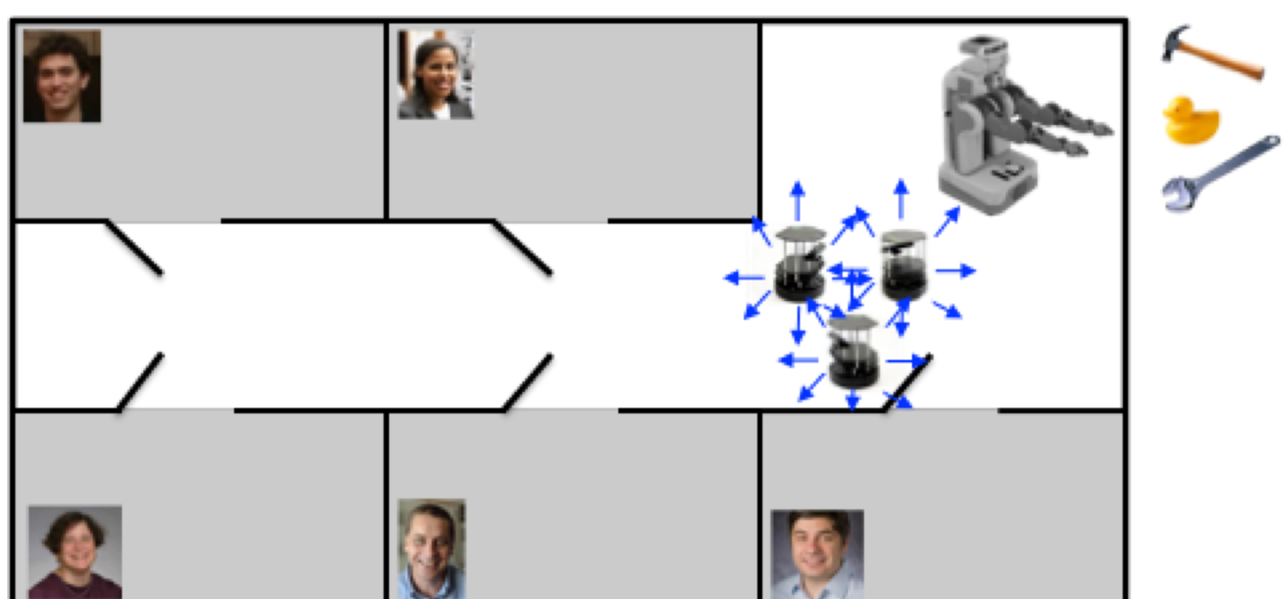
Our solutions (so far)

- Making limited approximations by:
    - Planning using hierarchy and sample-based methods
    - Learning solutions directly from data

Scaling up: macro-actions

Amato, Konidaris and Kaelbling - AAMAS 14

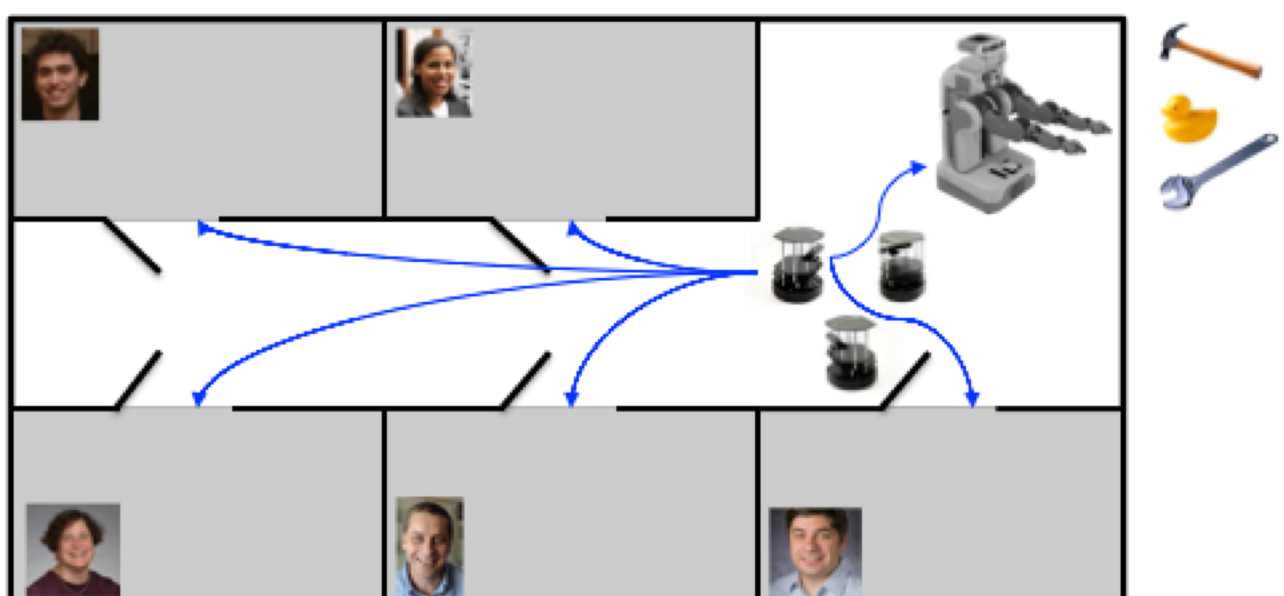
- Dec-POMDP methods model and solve at a low level (actions as control inputs)



Scaling up: macro-actions

Amato, Konidaris and Kaelbling - AAMAS 14

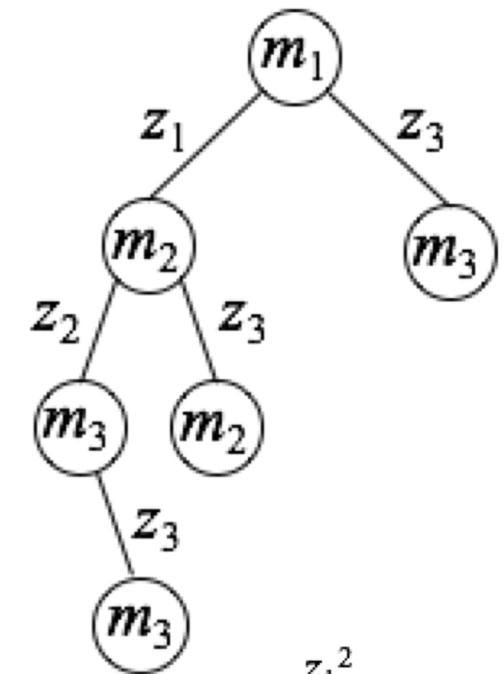
- Dec-POMDP methods model and solve at a low level (actions as control inputs)
- This is intractable (and unnecessary!) for real-world systems
- Often easy to plan for subgoals/subtasks
  - Set initial and terminal conditions (i.e., states)
  - Have expertly programmed controllers
- Allows for asynchronous decision-making
- Resulting model: MacDec-POMDP (macro-action Dec-POMDP)

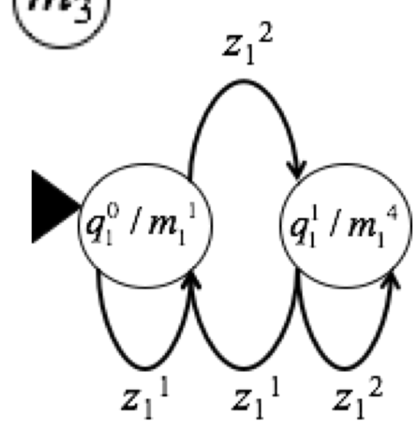


Macro-action solution representations

- Can extend policy representations to macro-action case
  - m = macro-action
  - z = high-level observation
- Finite-state controllers μ for each agent i defined with node set Q_i:
  - Action selection, λ: $Q_i \rightarrow M_i$
  - Node transitions, δ: $Q_i \times Z_i \rightarrow Q_i$
  - An initial node: $q_i^0 \in Q_i$
- But macro-actions finish at different times!
- Developed semi-Markov model, decentralized partially observable *semi-*Markov decision process (Dec-POSMDP)

Omidshafiei, Agha-mohammadi, Amato, Liu and How - IJRR 17

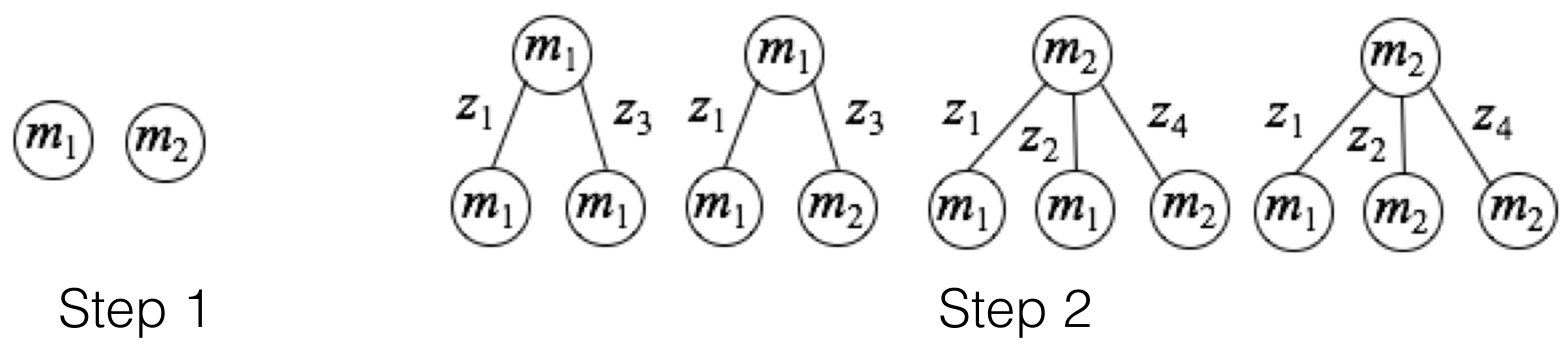




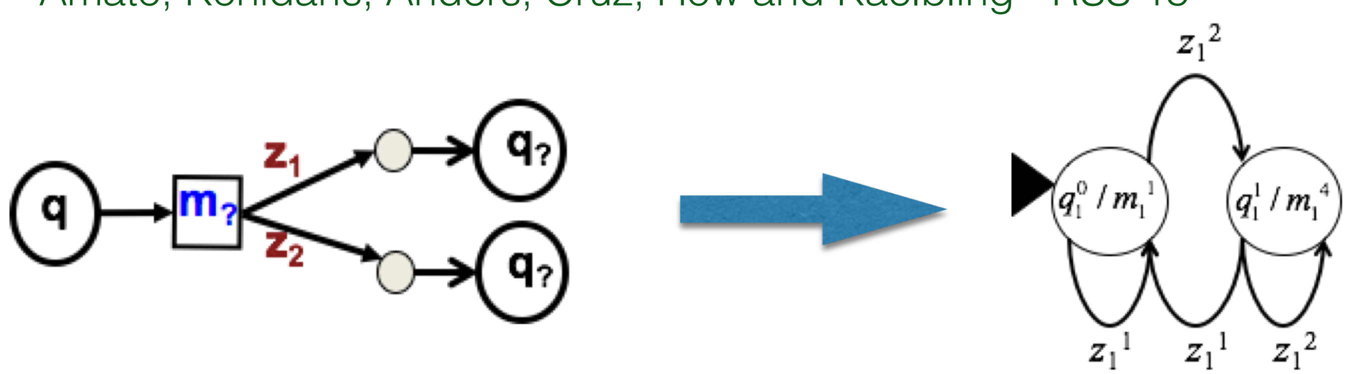
$$V^\mu(q,s) = R(s,\lambda(q)) + \sum_k^\infty \gamma^k \sum_{s',o} \Pr(s',k|s,\lambda(q)) \Pr(o|s',\lambda(q)) V^\mu(\delta(q,o),s')$$

Generating solutions

- Extended dynamic programming methods from Dec-POMDPs to build up trees
  Amato, Konidaris and Kaelbling - AAMAS 14



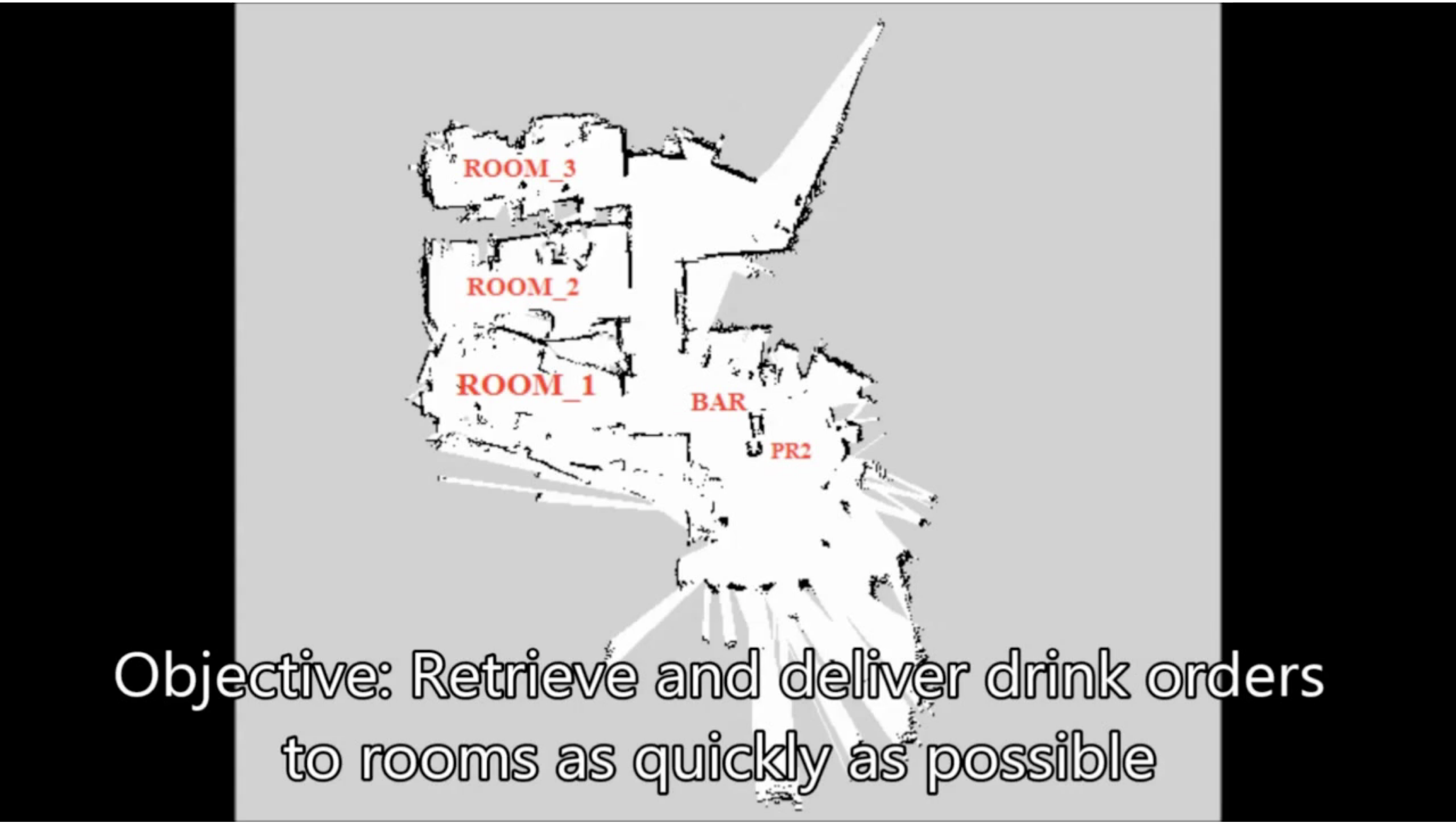
- Developed heuristic-based policy search methods to optimize controller parameters   Amato, Konidaris, Anders, Cruz, How and Kaelbling - RSS 15



- Perform evaluation using samples from simulator rather than have full model

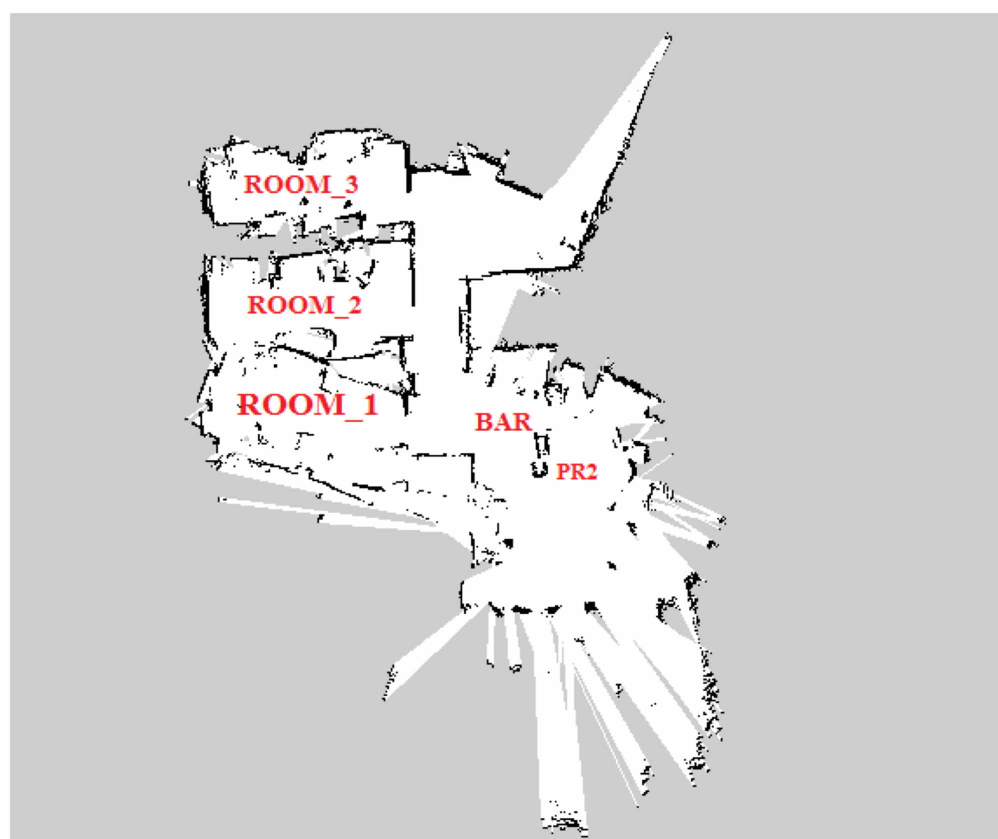
- Orders of magnitude faster and can solve problems that are orders of magnitude larger than previous Dec-POMDP methods (including robotics problems)

Cooperative beer delivery



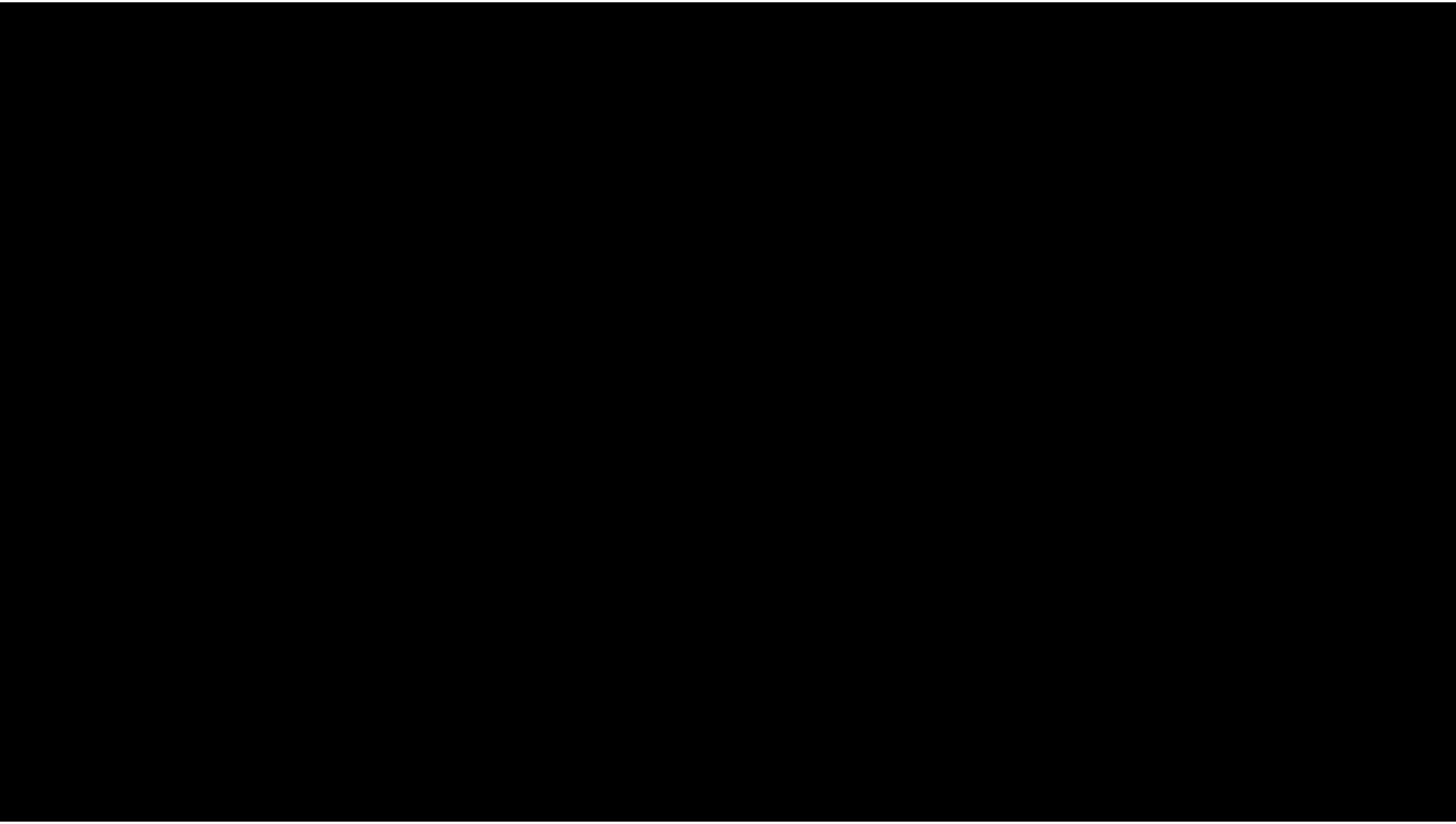
Cooperative beer delivery

- Controller values
  - 1-node controllers: 1231 (~13.98 drinks delivered)
  - 5-node controllers: 1296 (~14.56 drinks delivered)
  - Hand-coded solution: 917 (~11.13 drinks delivered)



Generating the ‘macro’-observations

Omidshafiei, Liu, Everett, Lopez, Amato, Liu, How, Vian - ICRA 17



Learning the solutions

- We may not have a model of the problem
- Want to learn solutions directly from data
- E.g., Learning a set of controllers from limited demonstrations
  Liu, Amato, Anesta, Griffith and How, AAAI 16

- Scalable to large state, macro-action and observation sets

Why can't we just use deep RL?

- Using deep RL for Dec-POMDPs has become a hot topic (e.g., Omidshafiei, Pazis, Amato, How and Vian, ICML 17, Foerster, Assael, de Freitas, and Whiteson, NIPS 16, Gupta, Egorov, Kochenderfer ICML 17)
- Some good results, but many open questions
  - Centralized vs. decentralized learning
  - Sample efficiency/online learning
  - Dealing with nonstationarity
  - Dealing with partial observability

Conclusions

- Dec-POMDPs represent a powerful probabilistic multi-agent framework
  - Considers outcome, sensor and communication uncertainty in a single framework
  - Can model any multi-agent coordination problem
  - Need to think about how to solve them
- Macro-actions provide an abstraction to improve scalability
- Learning methods can remove the need to generate a detailed multi-agent model
- Begun demonstrating scalability and quality in a number of domains, but a lot of great open questions to solve!
- A lot of great work by us and others as well (go see Frans Oliehoek!)

Postdoc(s) wanted!

- Come postdoc with me (or others) at Northeastern
  - Multi-agent RL
  - Deep RL
  - RL for robotic manipulation

- Come talk to or email me