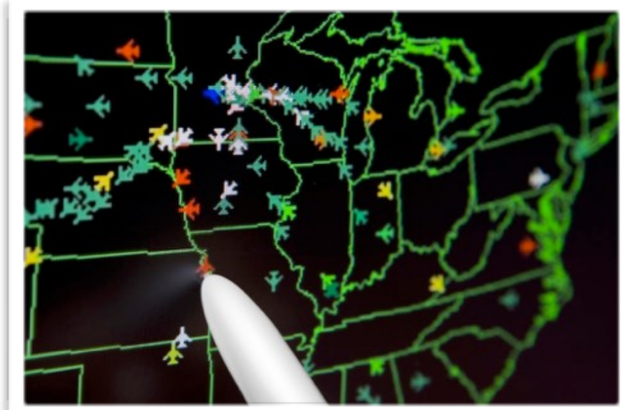
On Threshold Properties of the Optimal Policy for POMDPs on Partially Ordered Spaces

Erik Miehling — miehling@umich.edu
Department of Electrical & Computer Engineering
University of Michigan, Ann Arbor, MI, USA

FORCES - All Hands Meeting, Berkeley, CA — Aug 23-24, 2017

Motivation

- **Partially Observable Markov Decision Processes** (POMDPs) arise in many real-world settings where decisions need to be made over time and under uncertainty
- Finding an optimal policy is a notoriously difficult problem
- We aim to determine conditions such that the optimal policy has a nice form



FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS

What is a POMDP?

- State space, $\mathcal{S}$
- Action space, $\mathcal{U}$
- Transition probabilities, p_{ij}^{u}
- Observation space, $\mathcal{Y}$
- Observation probabilities, r_{jv}^{u}
- Cost function, $c : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$
- Discount factor, $\beta \in (0, 1)$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Solving the POMDP

- The sufficient information for making an optimal decision is summarized by a belief $\pi_t \in \Delta(\mathcal{S})$
- The goal is to find a policy $g : \Delta(\mathcal{S}) \to \mathcal{U}$ to minimize the total expected discounted cost

$$\mathbb{E}\left[\sum_{t=1}^{T} \beta^t c(s_t, u_t)\right]$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

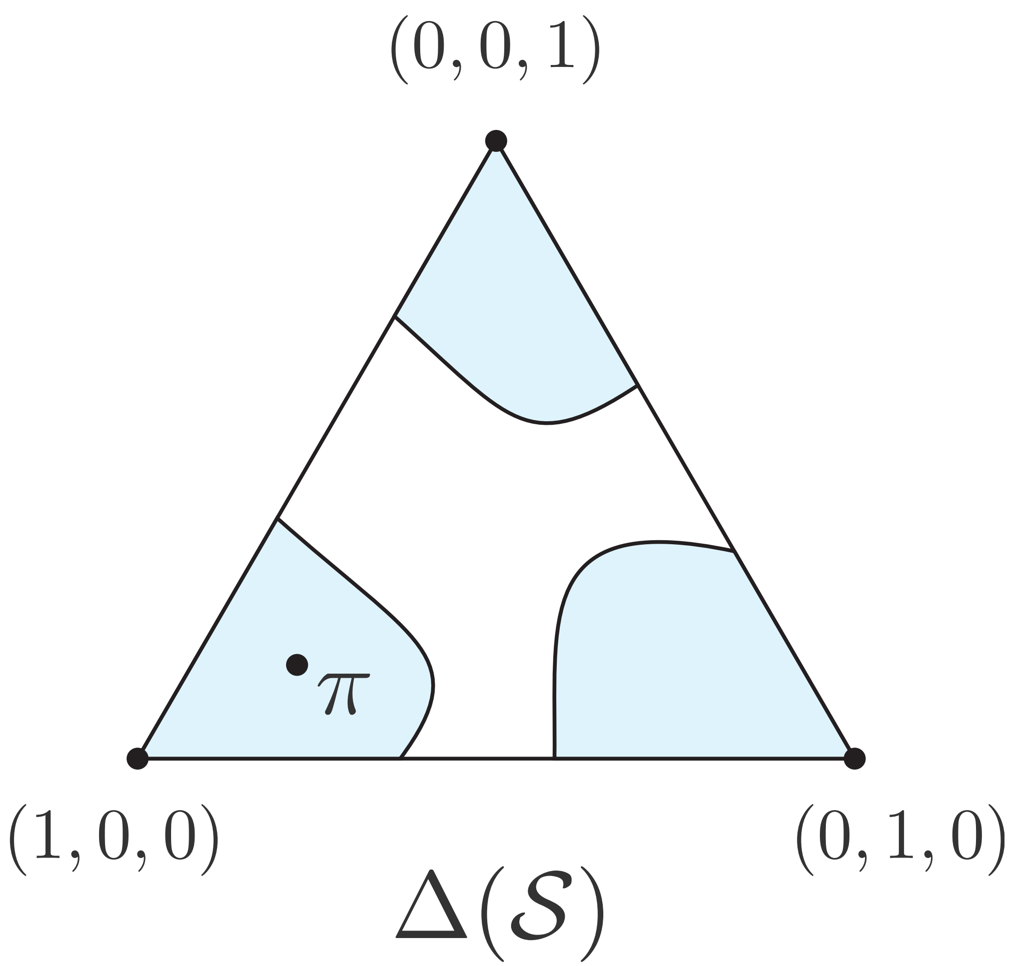
Solving the POMDP

- The sufficient information for making an optimal de
- Th ize the

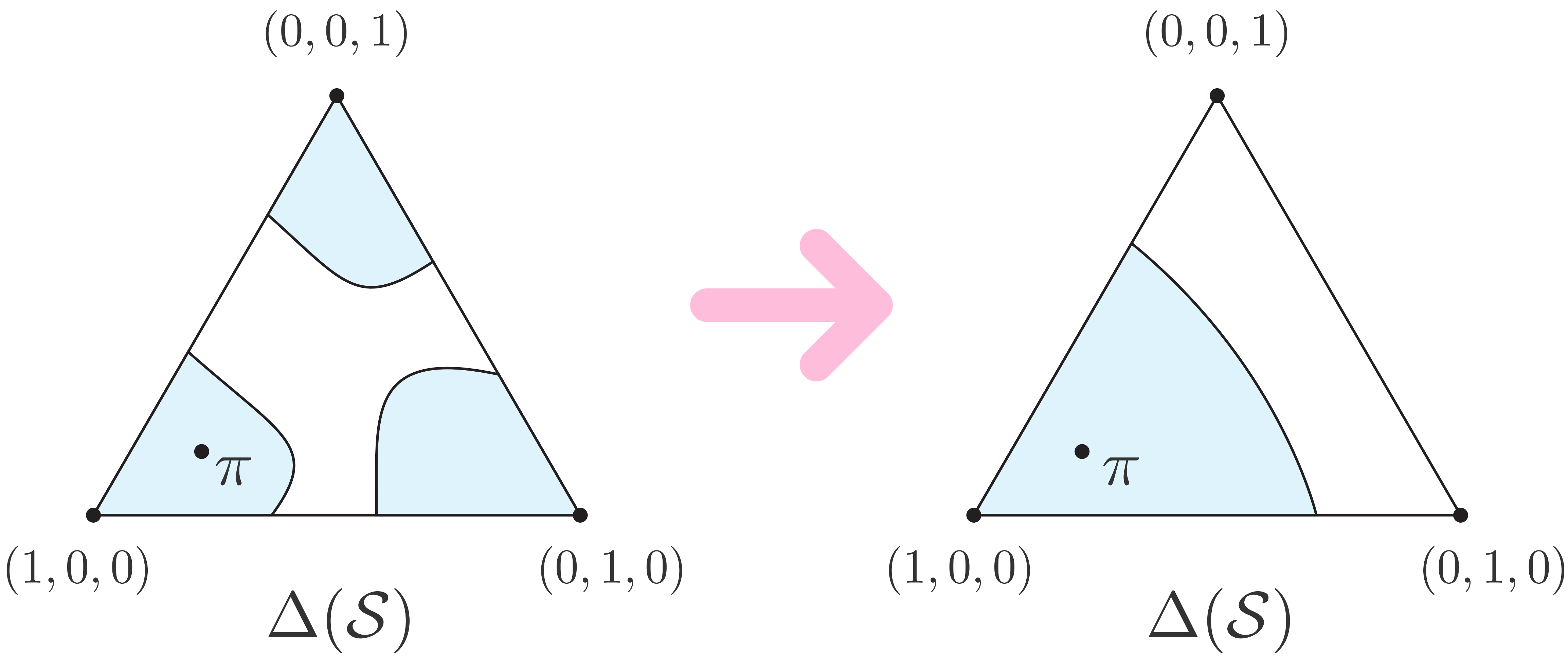
Under what conditions does the optimal policy have desirable structure?

$$\mathbb{E}\left[\sum_{t=1}^{T} \beta^t c(s_t, u_t)\right]$$

Structured Policies



Structured Policies



First Order Stochastic Dominance

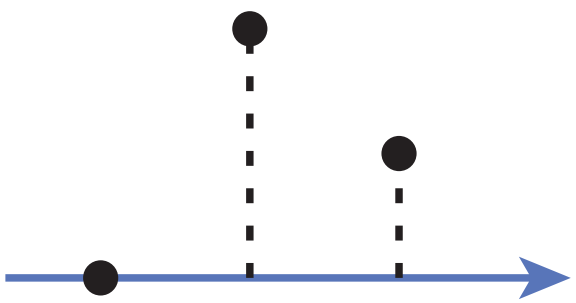
Definition (First Order Stochastic Dominance)

Given elements $\pi, \pi' \in \Delta(\mathcal{S})$, π is said to dominate π' with respect to first order stochastic dominance (FOSD), written $\pi \succeq_{st} \pi'$, if

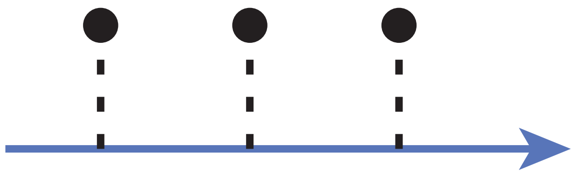
$$\sum_{j=i}^{n} \pi_j \geq \sum_{j=i}^{n} \pi'_j$$

for all $i = 1, \ldots, n$.

FOSD is Fragile

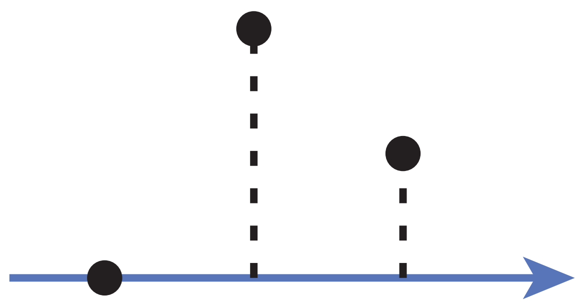


$$\pi = \left(0, \frac{2}{3}, \frac{1}{3}\right)$$

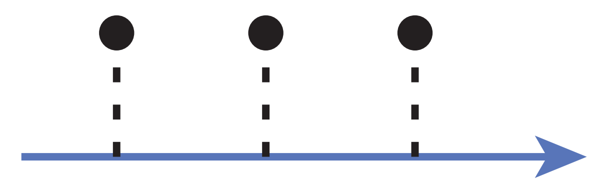


$$\pi' = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

FOSD is Fragile

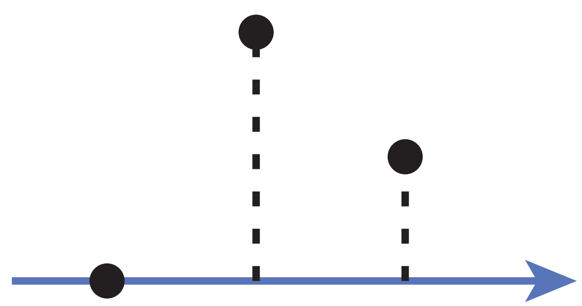


$$\succeq_{st}$$

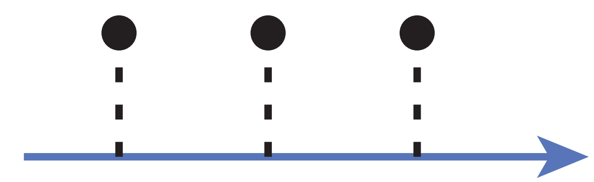


$$\pi = \left(0, \frac{2}{3}, \frac{1}{3}\right) \qquad \pi' = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

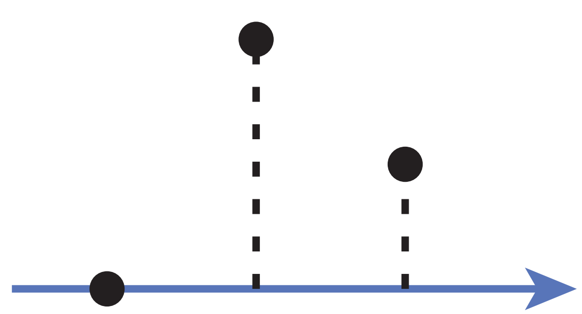
FOSD is Fragile

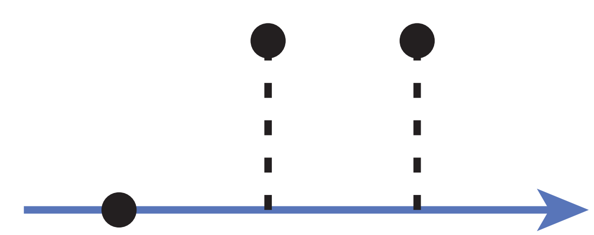


$$\succeq_{st}$$



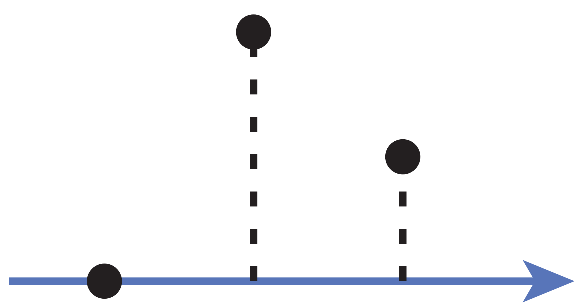
$$\pi = \left(0, \frac{2}{3}, \frac{1}{3}\right) \qquad \pi' = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$



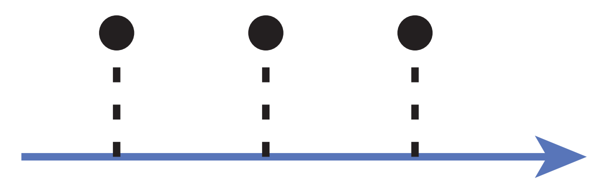


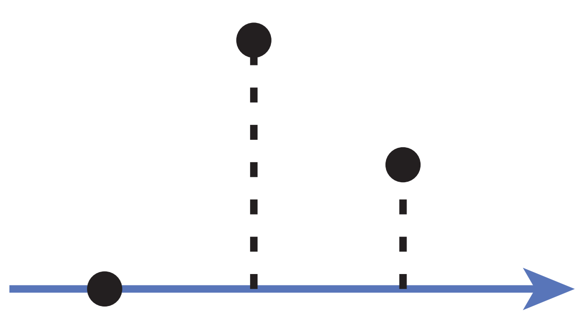
$$\pi = \left(0, \frac{2}{3}, \frac{1}{3}\right) \qquad \pi' = \left(0, \frac{1}{2}, \frac{1}{2}\right)$$

FOSD is Fragile

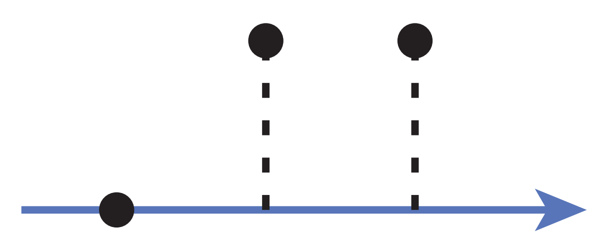


$$\pi = \left(0, \frac{2}{3}, \frac{1}{3}\right) \quad \succeq_{st} \quad \pi' = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$





$$\pi = \left(0, \frac{2}{3}, \frac{1}{3}\right) \quad \preceq_{st} \quad \pi' = \left(0, \frac{1}{2}, \frac{1}{2}\right)$$



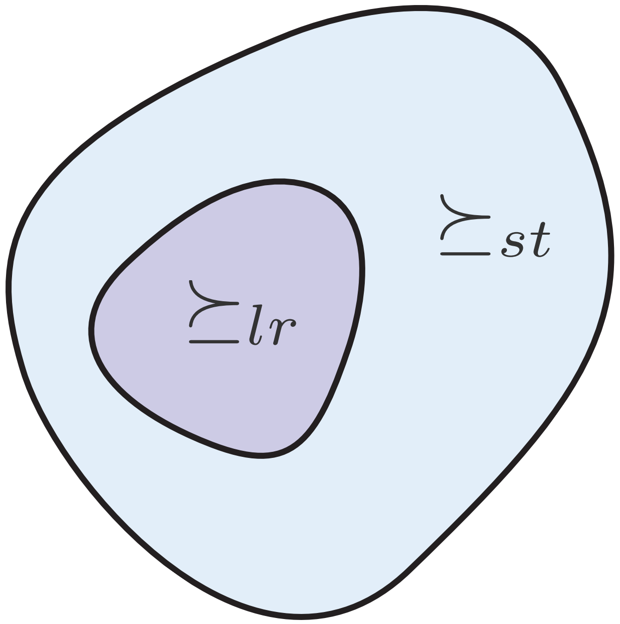
Monotone Likelihood Ratio Order

Definition (Monotone Likelihood Ratio)

Given elements $\pi, \pi' \in \Delta(\mathcal{S})$, π is said to be greater than π' with respect to the monotone likelihood ratio (MLR), written $\pi \succeq_{lr} \pi'$, if

$$\pi_i \pi'_j \geq \pi_j \pi'_i$$

for all $i \geq j$.



[Lovejoy 1987, **Some Monotonicity Results for Partially Observed Markov Decision Processes**]

Partial Orders

- The existing orders assumed that the underlying space is totally ordered
- We are interested in spaces $\mathcal{S}$ that are partially ordered by $\succeq$, ***i.e.*** posets $(\mathcal{S}, \succeq)$
- That is, for some states $s, s' \in \mathcal{S}$, neither $s \succeq s'$ nor $s' \succeq s$ hold, such cases are denoted by $s \parallel s'$
- **Example:** under the element-wise partial order $\succeq_e$

$$(2,1) \succeq_e (1,1) \qquad\qquad (1,2) \parallel (2,1)$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Generalized FOSD Order

Definition (Generalized FOSD, White 1979)

Given elements $\pi, \pi' \in \Delta(\mathcal{S})$, π is said to dominate π' with respect to generalized first order stochastic dominance (GFOSD), written $\pi \succeq_{gst} \pi'$, if

$$\pi I_K \geq \pi' I_K$$

for all $K \in \mathcal{K} = \{K \subseteq \mathcal{S} \mid s_j \in K, s_i \succeq s_j \implies s_i \in K\}$.

GFOSD Example

- Consider the state-space $\mathcal{S} = \{s_1, s_2, s_3\}$ and partial order $\succeq$ such that

$$s_3 \succeq s_1$$
$$s_3 \succeq s_2$$
$$s_1 \parallel s_2$$

- We have $\pi \succeq_{gst} \pi'$ if and only if

$$\pi_1 + \pi_3 \geq \pi'_1 + \pi'_3$$
$$\pi_2 + \pi_3 \geq \pi'_2 + \pi'_3$$
$$\pi_3 \geq \pi'_3$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Existing Work

	total order	partial order
FOSD	**Porteus 1975**	**White 1979**
MLR	**Lovejoy 1987**	**???**

The POMDP Model

- State space, $\mathcal{S}$
- Action space, $\mathcal{U}$
- Transition probabilities, p_{ij}^u
- Observation space, $\mathcal{Y}$
- Observation probabilities, r_{jv}^u
- Cost function, $c : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$
- Discount factor, $\beta \in (0, 1)$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

The POMDP Model

- State space, $\mathcal{S}$ $\longleftarrow$ $(\mathcal{S}, \succeq)$
- Action space, $\mathcal{U}$ $\longleftarrow$ $\mathcal{U} = \{u_0(\text{null}), u_1(\text{reset})\},\ u_1 \geq u_0$
- Transition probabilities, p_{ij}^u
- Observation space, $\mathcal{Y}$ $\longleftarrow$ $(\mathcal{Y}, \succeq_{\mathcal{Y}})$
- Observation probabilities, r_{jv}^u
- Cost function, $c : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$
- Discount factor, $\beta \in (0, 1)$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

The POMDP Model

- State space, $\mathcal{S}$ $\longleftarrow$ $(\mathcal{S}, \succeq)$
- Action space, $\mathcal{U}$ $\longleftarrow$ $\mathcal{U} = \{u_0(\text{null}), u_1(\text{reset})\},\ u_1 \geq u_0$
- Transition probabilities, p_{ij}^u
- Observation space, $\mathcal{Y}$ $\longleftarrow$ $(\mathcal{Y}, \succeq_{\mathcal{Y}})$
- Observation probabilities, r_{jv}^u
- Cost function, $c : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$
- Discount factor, $\beta \in (0, 1)$

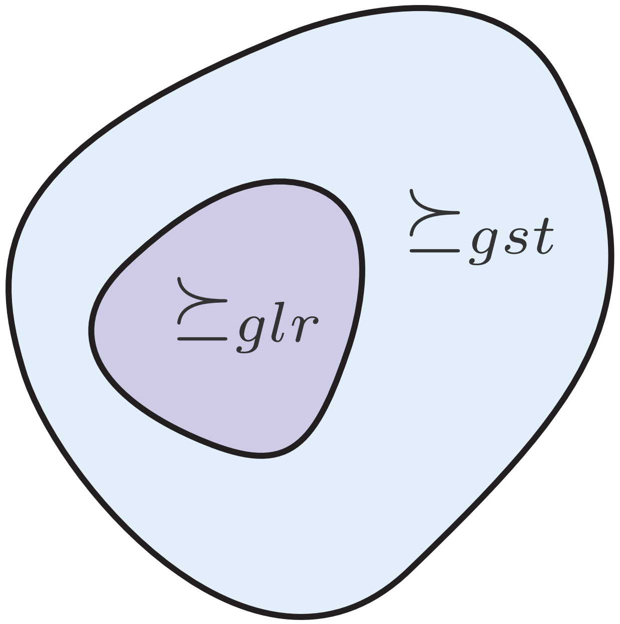
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

GMLR Order

Definition (Generalized Monotone Likelihood Ratio)

Given elements $\pi, \pi' \in \Delta(\mathcal{S})$, π is said to be greater than π' with respect to the generalized monotone likelihood ratio (GMLR), written $\pi \succeq_{glr} \pi'$, if

$$\begin{aligned} \pi_i \pi'_j &\geq \pi_j \pi'_i \quad \text{for } s_i \succeq s_j \\ \pi_i \pi'_j &= \pi_j \pi'_i \quad \text{for } s_i \parallel s_j \end{aligned}$$



GMLR Example

- Recall the partially ordered state-space $\mathcal{S} = \{s_1, s_2, s_3\}$ from before where $s_3 \succeq s_1$, $s_3 \succeq s_2$, and $s_1 \parallel s_2$

- Two comparable beliefs under $\succeq_{glr}$

$$\pi = (0.2, 0.1, 0.7) \qquad \succeq_{glr} \quad \pi' = (0.4, 0.2, 0.4)$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

GMLR Example

- Recall the partially ordered state-space $\mathcal{S} = \{s_1, s_2, s_3\}$ from before where $s_3 \succeq s_1$, $s_3 \succeq s_2$, and $s_1 \parallel s_2$

- Two comparable beliefs under $\succeq_{glr}$

$$\pi = (0.2, 0.1, 0.7) \qquad \succeq_{glr} \qquad \pi' = (0.4, 0.2, 0.4)$$

- The following beliefs are not comparable under $\succeq_{glr}$

$$\pi = (1, 0, 0)$$
$$\pi' = (0, 1, 0)$$

$$\underset{1}{\pi_1}\,\underset{1}{\pi'_2} \not= \underset{0}{\pi_2}\,\underset{0}{\pi'_1}$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

GMLR-order Preserving Matrices

Definition (GTP$_2$ Matrix)

A stochastic matrix is termed generalized totally positive of order 2 (GTP$_2$) if for all $s_k \succeq s_l$

$$p_{lj}p_{ki} - p_{kj}p_{li} \geq 0 \quad \text{for } s_i \succeq s_j$$

$$p_{lj}p_{ki} - p_{kj}p_{li} = 0 \quad \text{for } s_i \parallel s_j$$

Example

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 1 \end{bmatrix}$$

$$s_3 \succeq s_1$$

$$s_3 \succeq s_2$$

$$s_1 \parallel s_2$$

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

GMLR-order Preserving Matrices

Definition (GTP$_2$ Matrix)

A stochastic matrix is termed generalized totally positive of order 2 (GTP$_2$) if for all $s_k \succeq s_l$

$$p_{lj}p_{ki} - p_{kj}p_{li} \geq 0 \quad \text{for } s_i \succeq s_j$$

$$p_{lj}p_{ki} - p_{kj}p_{li} = 0 \quad \text{for } s_i \parallel s_j$$

Example

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 1 \end{bmatrix}$$

$$s_3 \succeq s_1$$

$$s_3 \succeq s_2$$

$$s_1 \parallel s_2$$

Proposition

If $\pi \succeq_{glr} \pi'$ then $\pi P \succeq_{glr} \pi' P$ if and only if P is GTP$_2$.

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Conditions for Threshold Policy

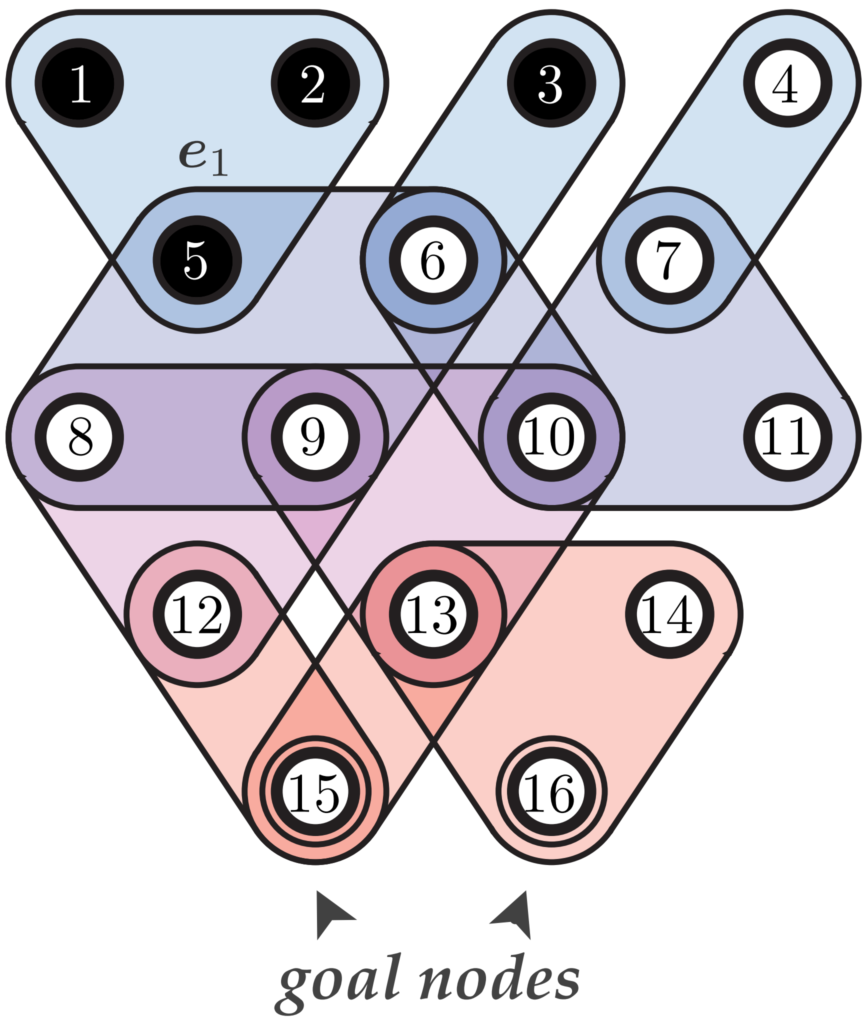
Let $\pi \succeq_{glr} \pi'$ and assume that

1. $c(s, u)$ is increasing in s on $(\mathcal{S}, \succeq)$
2. $c(s, u_1) - c(s, u_0)$ is decreasing in s on $(\mathcal{S}, \succeq)$
3. P^u is GTP$_2$ for each $u \in \mathcal{U}$
4. $r_{iv} r_{jw} = r_{jv} r_{iw}$ for any $s_i \parallel s_j$ in $\mathcal{S}$, $y_v \succeq_{\mathcal{Y}} y_w$ in $\mathcal{Y}$
   or $s_i \succeq s_j$ in $\mathcal{S}$, $y_v \parallel_{\mathcal{Y}} y_w$ in $\mathcal{Y}$
5. $r_i \succeq_{glr} r_j$ for all $s_i \succeq s_j$

then $g_t^*(\pi) \geq g_t^*(\pi')$ for any t.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

An Application in Security



- **Question**: at what point should the network be reset?

$$(\mathcal{S}, \succeq) = (\mathcal{S}, \supseteq)$$
$$(\mathcal{Y}, \succeq_{\mathcal{Y}}) = (2^n, \supseteq)$$

- Transition matrix is GTP_2
- Reasonable conditions on observation process
  - Optimal policy is threshold

Summary

- We have derived conditions to ensure that the optimal policy takes a threshold form
- The results are applicable to any partially observable domain with a binary action space, $\mathcal{U} = \{u_0, u_1\}$

  u_0 : lets system continue uninterrupted

  u_1 : resets system back to the initial state with certainty

- Can be exploited computationally to design efficient algorithms

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Acknowledgments

- Special thanks to the following funding sources

- **NSF —** Foundations Of Resilient CybEr-physical Systems (FORCES)

Grant: **CNS-1238962**

- **ARO MURI —** Adversarial and Uncertain Reasoning for Adaptive Cyber Defense: Building the Scientific Foundations

Grant: **W911NF-13-1-0421**

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Thank You

Questions?

Contact

umich.edu/~miehling/

miehling@umich.edu

FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS