Cooperative Systems Design

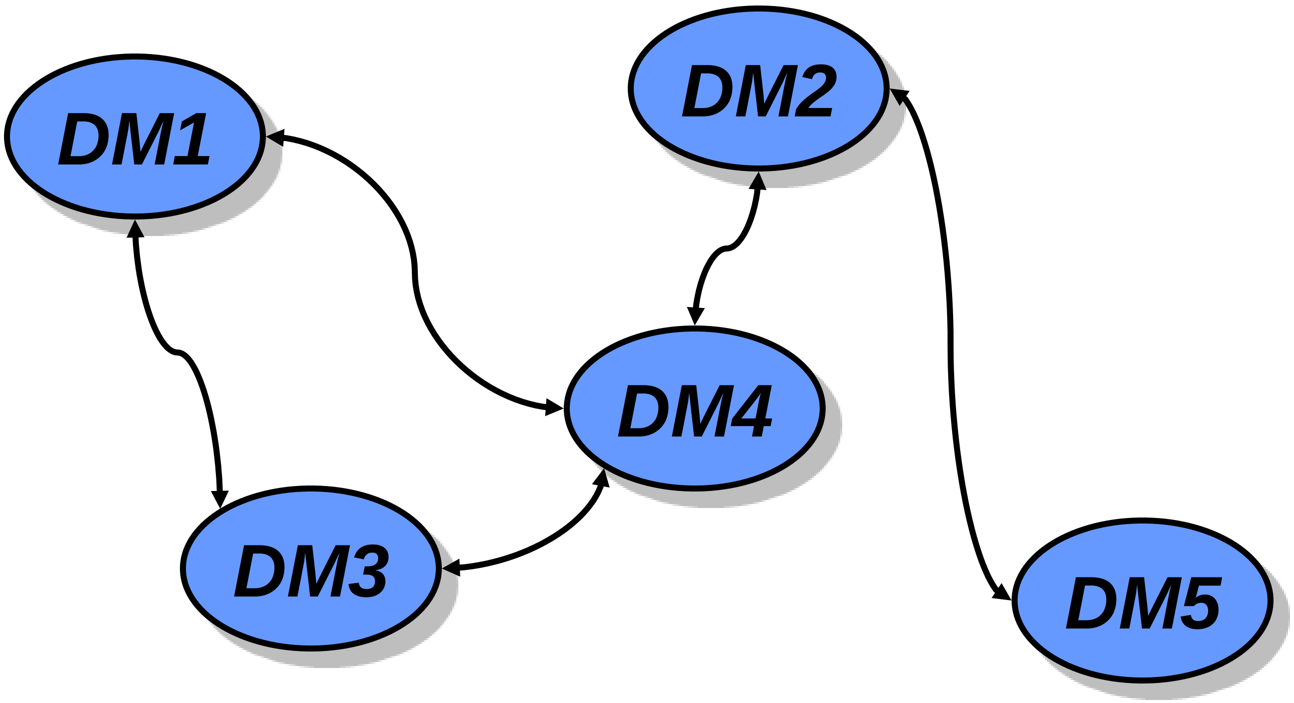
Gürdal Arslan

Electrical Engineering
University of Hawaii

www2.hawaii.edu/~gurdal/

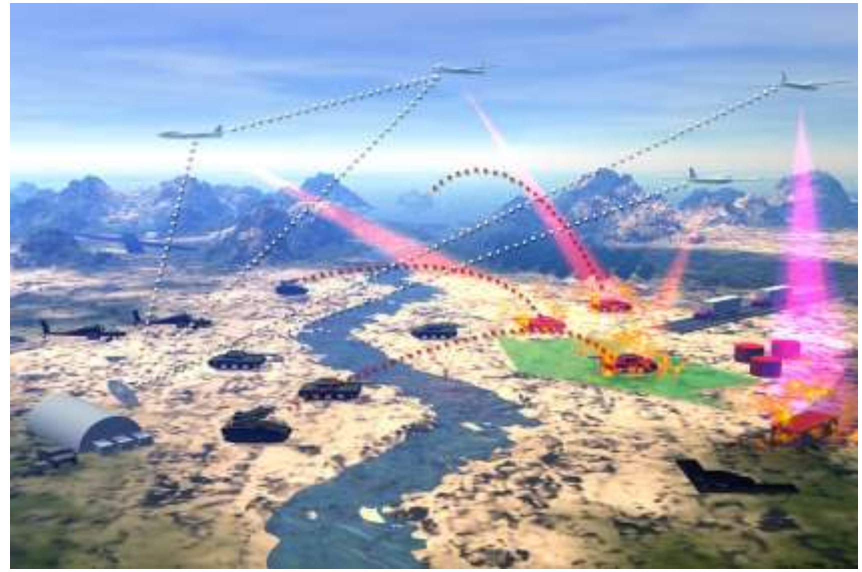
National Science Foundation & Lockheed Martin

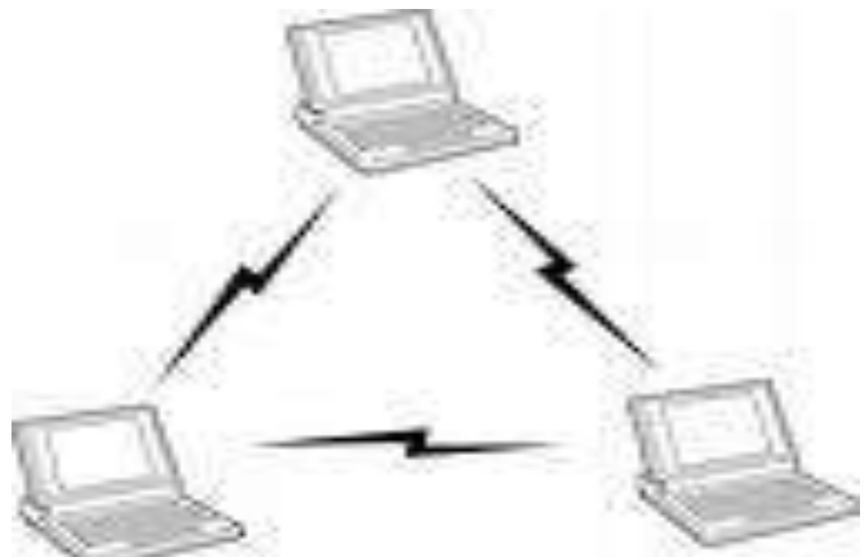
Cooperative Systems



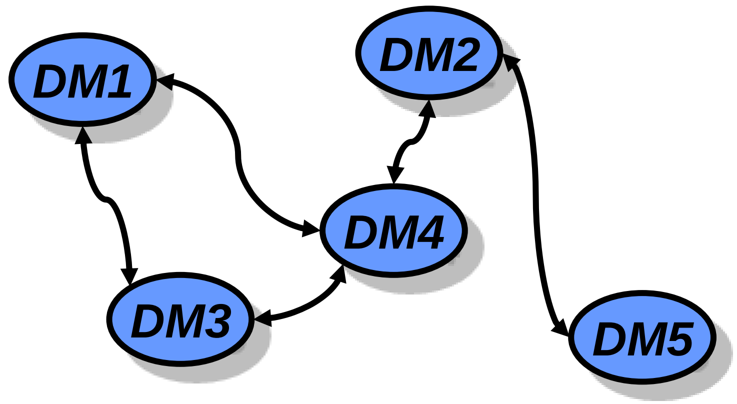
- **Decision Makers (DMs) are *self-interested***
- **Individual rewards depends on collective actions**
- **Collective behavior depends on individual actions**

Cooperative Systems: Natural and Virtual



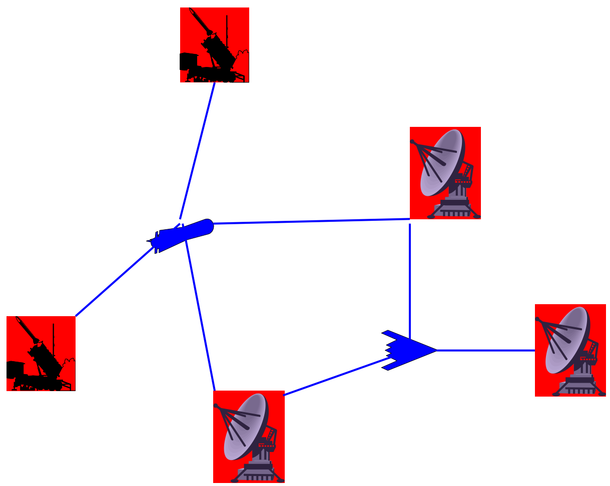


Cooperative Systems Design



- **Optimize a global objective through selfish DMs**
- **Design Problem:**
  - **Utility Design (tell DMs what to optimize)**
  - **Negotiation Mechanism Design (tell DMs how to optimize)**

Overview



Autonomous Target Assignment

(MURI)

Cooperative Systems Design:

A Markov Games Approach

(NSF)

Autonomous Sensor Allocation

for Submarine Tracking

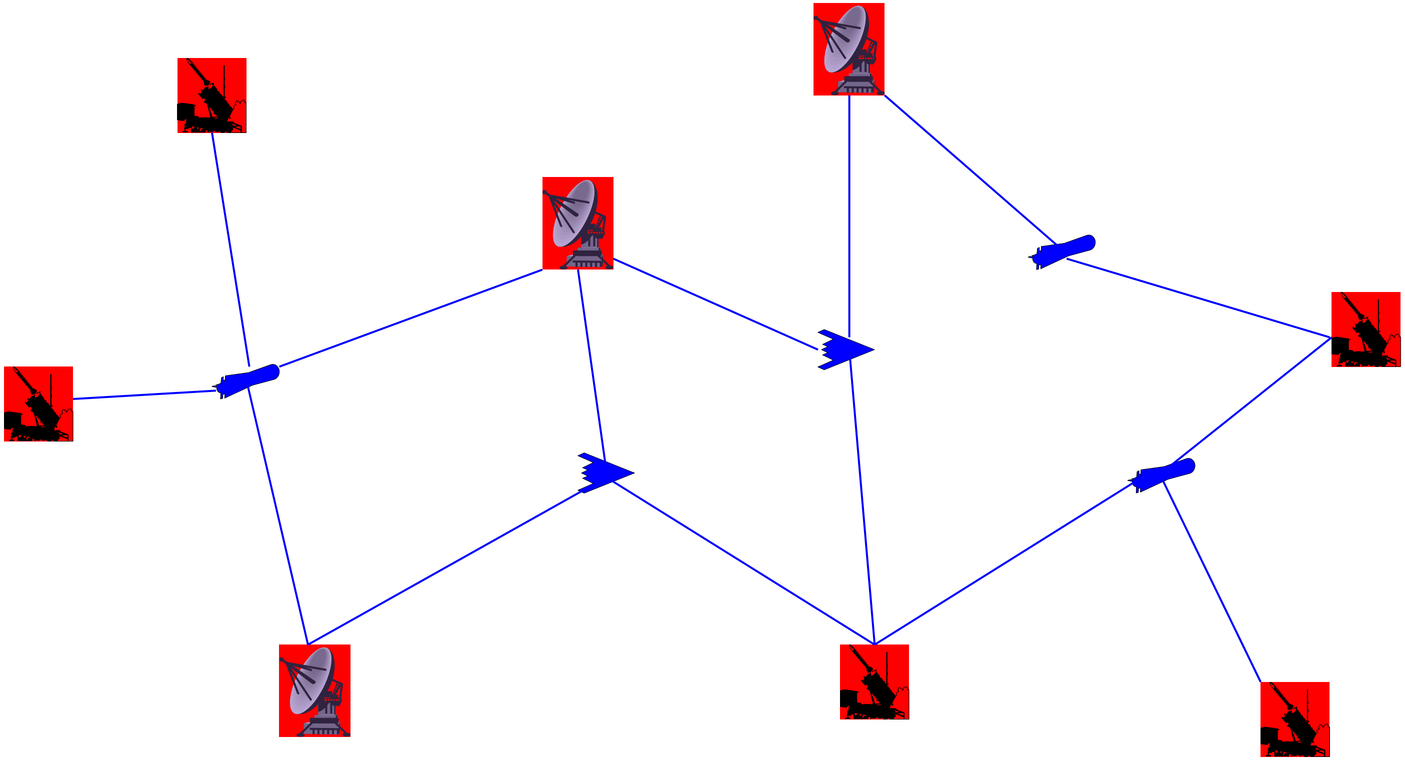
(Lockheed Martin)

Negotiated Signaling and Power Management

in Wireless MIMO Interference Systems

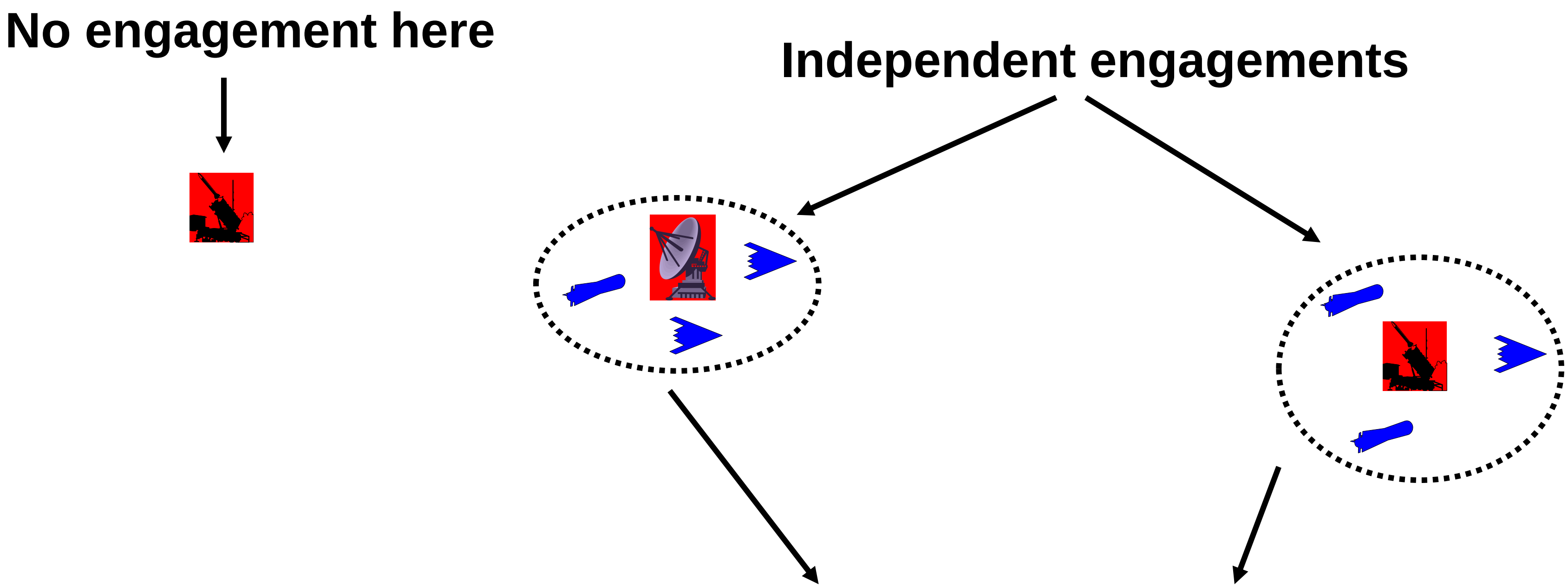
(UH)

Setup for Target Assignment Problem



max   Global Utility (assignment)

Global Utility



$$\text{Global Utility} = \sum_j \underbrace{\text{Utility generated at target j}}_{}$$

$\downarrow$ (for example)

$$E\,[\text{ total value of (destroyed target} - \text{ vehicles lost) }]$$

Joint Optimization

Global Utility

Assignment Profile : $(a_1, \ldots, a_n)$

$$\max_{a} \; U_g(a)$$

- **Can be formulated as an integer programming problem**

  - Computationally hard

  - Relaxation techniques available for suboptimal solutions

- **Decentralized implementation**

  - Requires global information

  - Agreement issues can arise

Game Theory Formulation

- Vehicles are self-interested players with private utilities

$$U_i(a)$$

- A vehicle need not know other vehicles' utilities.
- Individual utilities depend on local information only.
- Vehicles negotiate an agreeable assignment.

Autonomous Target Assignment Problem

Design

- **Vehicle utilities** $U_i(a)$
- **Negotiation mechanisms**

so that

vehicles agree on an assignment with high Global Utility

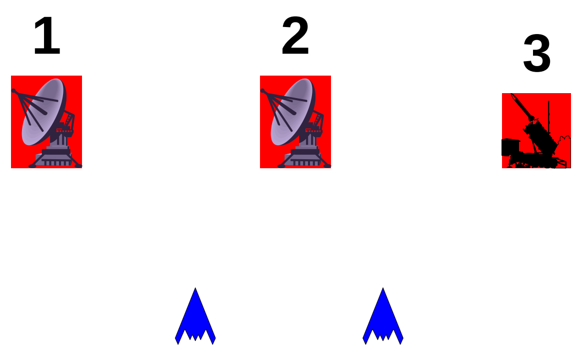
using

- **low computational power**
- **low inter-vehicle communication**

Agreeable Assignment - Nash Equilibrium

- An assignment is a (pure) Nash equilibrium
  if no player has an incentive to unilaterally deviate from it.

- Example :



	1	2	3
1	2,1	0,0	0,0
2	0,0	1,2	0,0
3	0,0	0,0	3,3

- Pure Nash equilibria : (1,1), (2,2), (3,3)

- Mixed Nash equilibria : ([.54 .27 .18] , [.27 .54 .18])

Utility Design

- **Vehicle utilities should be aligned with Global Utility**
- **Ideal alignment :**
  - **Only globally optimal assignments should be agreeable**
  - **Not possible without computing globally optimal assignments**
- **Relaxed alignment (factoredness in Wolpert et al. 2000) :**

$$U_i(a_1,\ldots,\tilde{a}_i,\ldots,a_n) > U_i(a_1,\ldots,a_i,\ldots,a_n)$$

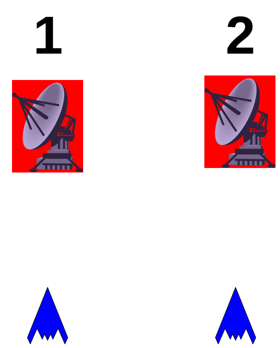
$$\Leftrightarrow$$

$$U_g(a_1,\ldots,\tilde{a}_i,\ldots,a_n) > U_g(a_1,\ldots,a_i,\ldots,a_n)$$

  - **Globally optimum assignment is always agreeable (pure Nash)**

Aligned Utilities - Team Play

- For every vehicle, $$U_i(a) = U_g(a)$$
- Example :



	1	2
1	2 , 2	0 , 0
2	0 , 0	1 , 1

Suboptimal Nash

- Not localized
  - Each vehicle needs global information
  - Low Signal-to-Noise-Ratio (Wolpert et al. 2000)

Aligned Utilities - Wonderful Life Utility (Wolpert et al. 2000)

- **Marginal contribution of vehicle # i to Global Utility, i.e.,**

$$U_i(a) = U_g(a) - U_g(a : \underbrace{a_i = "0"}_{\text{no engagement}})$$

- **Localized :**

  - Equal marginal contribution to engagements within range

  - Signal-to-Noise-Ratio is maximized

Aligned Utilities - Wonderful Life Utility (Wolpert et al. 2000)

- **Aligned :**

$$U_i(a_1,\ldots,\tilde{a}_i,\ldots,a_n) - U_i(a_1,\ldots,a_i,\ldots,a_n)$$

$$=$$

$$U_g(a_1,\ldots,\tilde{a}_i,\ldots,a_n) - U_g(a_1,\ldots,a_i,\ldots,a_n)$$

- **Leads to a Potential Game with potential** $U_g(a)$
- **Convergent negotiation mechanisms for potential games**

A Misaligned Utility Structure

- **Equally Shared Utilities :**

$$U_i(a) = \frac{\text{utility generated by engagement with } a_i}{\text{number of participating vehicles}}$$

- **Hence**

$$U_g(a) = \sum_i U_i(a)$$

- **Global optimum may not be Nash agreeable**
- **A pure Nash agreeable assignment may not exists at all !**

Negotiation Mechanisms

- At step k, vehicle # i proposes a target

$$a_i(k)$$

based on the past proposal profiles

$$a(1),\ldots,a(k-1)$$

- Is there a reasonable negotiation mechanism that leads to a Nash equilibrium ?
- Adopt learning methods in repeated games

Spatial Adaptive Play

- **At each negotiation step, only one randomly chosen vehicle updates its proposal**

- **Updating vehicle proposes** a_i **at step k with probability**

$$p_i^{a_i}(k) := P\ a_i(k) = a_i$$

which maximizes

$$p_i(k) = \arg\max_{p_i} \ \underset{a_i \sim p_i}{E}\ U_i(a_i, a_{-i}(k-1)) + \tau H(p_i)$$

Spatial Adaptive Play

- $p_i(k)$ is given by Gibbs distribution

$$p_i^{a_i}(k) = \frac{e^{U_i(a_i, a_{-i}(k-1))/\tau}}{\sum_{a_i} e^{U_i(a_i, a_{-i}(k-1))/\tau}}$$

- For potential games, SAP induces a Markov Chain with

$$\lim_{k} P\; a(k) = a = \frac{e^{U_g(a)/\tau}}{\sum_{a} e^{U_g(a)/\tau}}$$

Spatial Adaptive Play

- As $\tau \downarrow 0$, we have

$$\frac{e^{U_g(a)/\tau}}{\sum_a e^{U_g(a)/\tau}} \to 1 \quad \text{for } a \in \arg\max U_g(a)$$

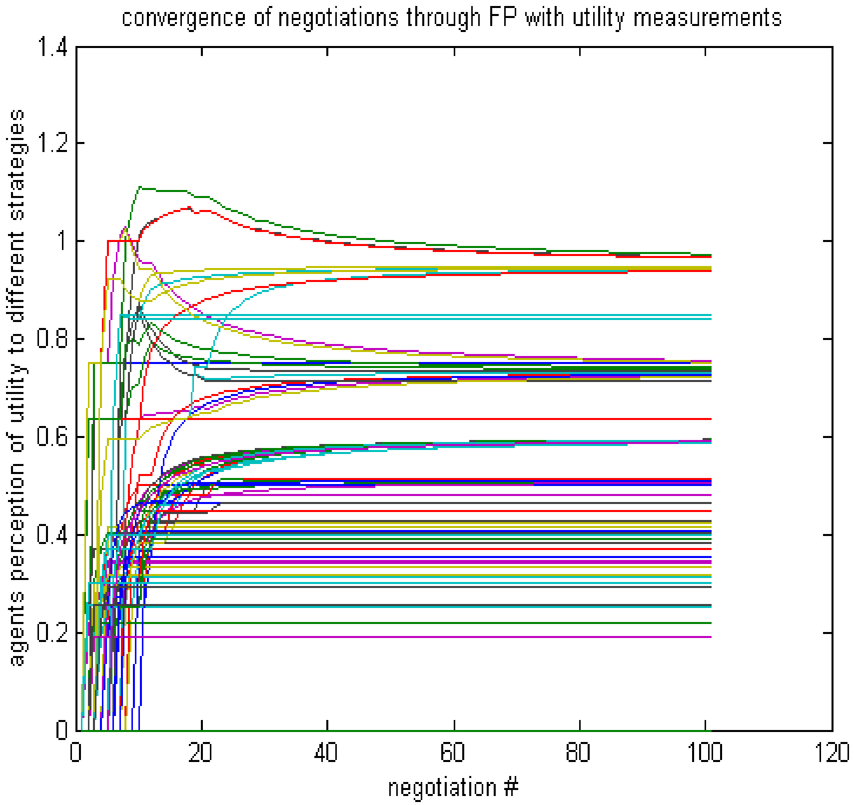
- Therefore,

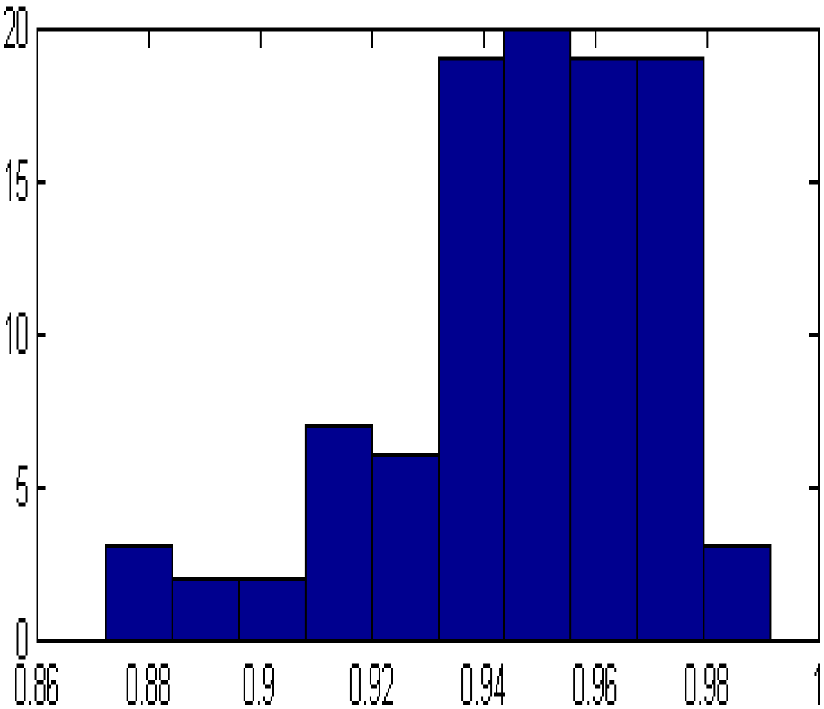
$$\lim_{\tau \downarrow 0} P\left(a(\infty) \in \arg\max U_g(a)\right) = 1$$

Near Optimum Performance

- **Example:**

40 uniform weapons negotiate 40 non-uniform targets

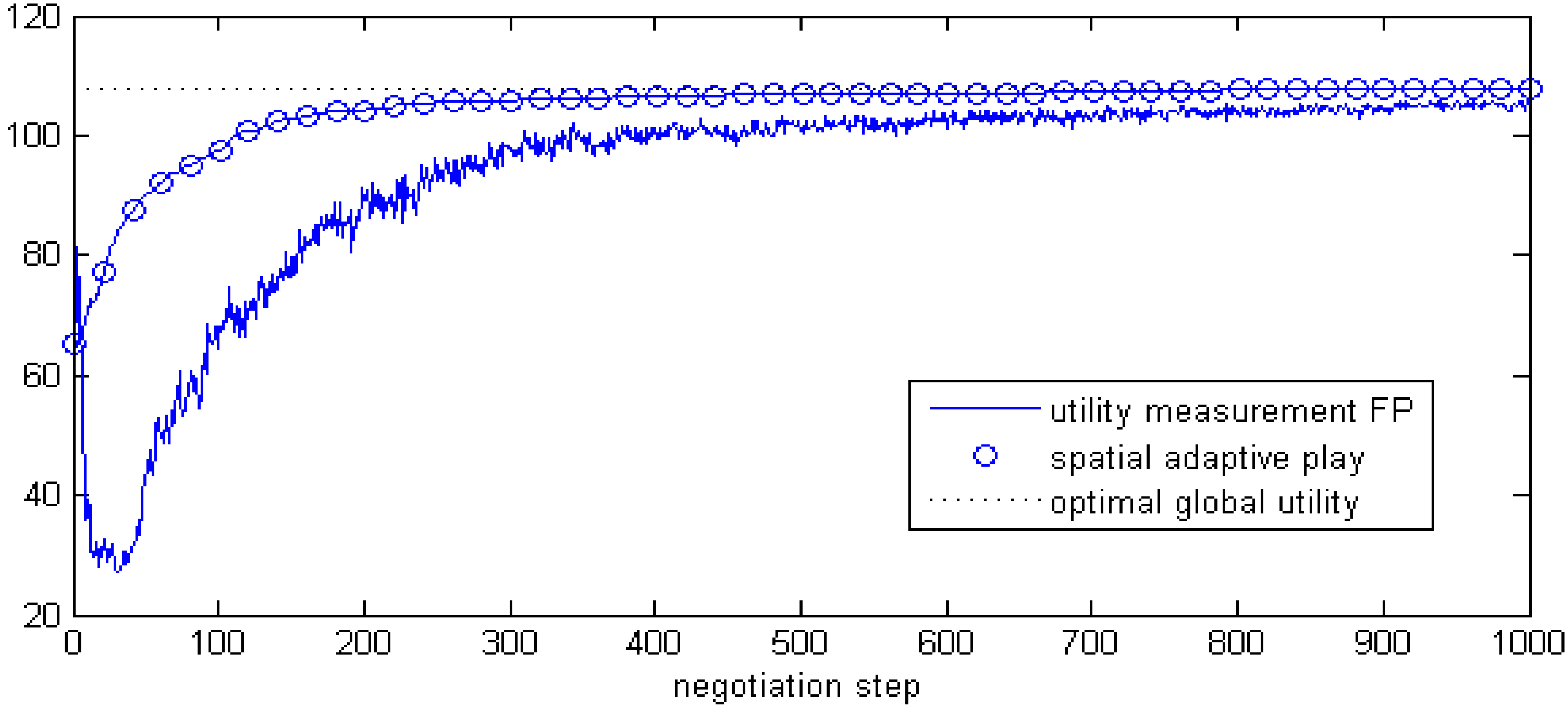




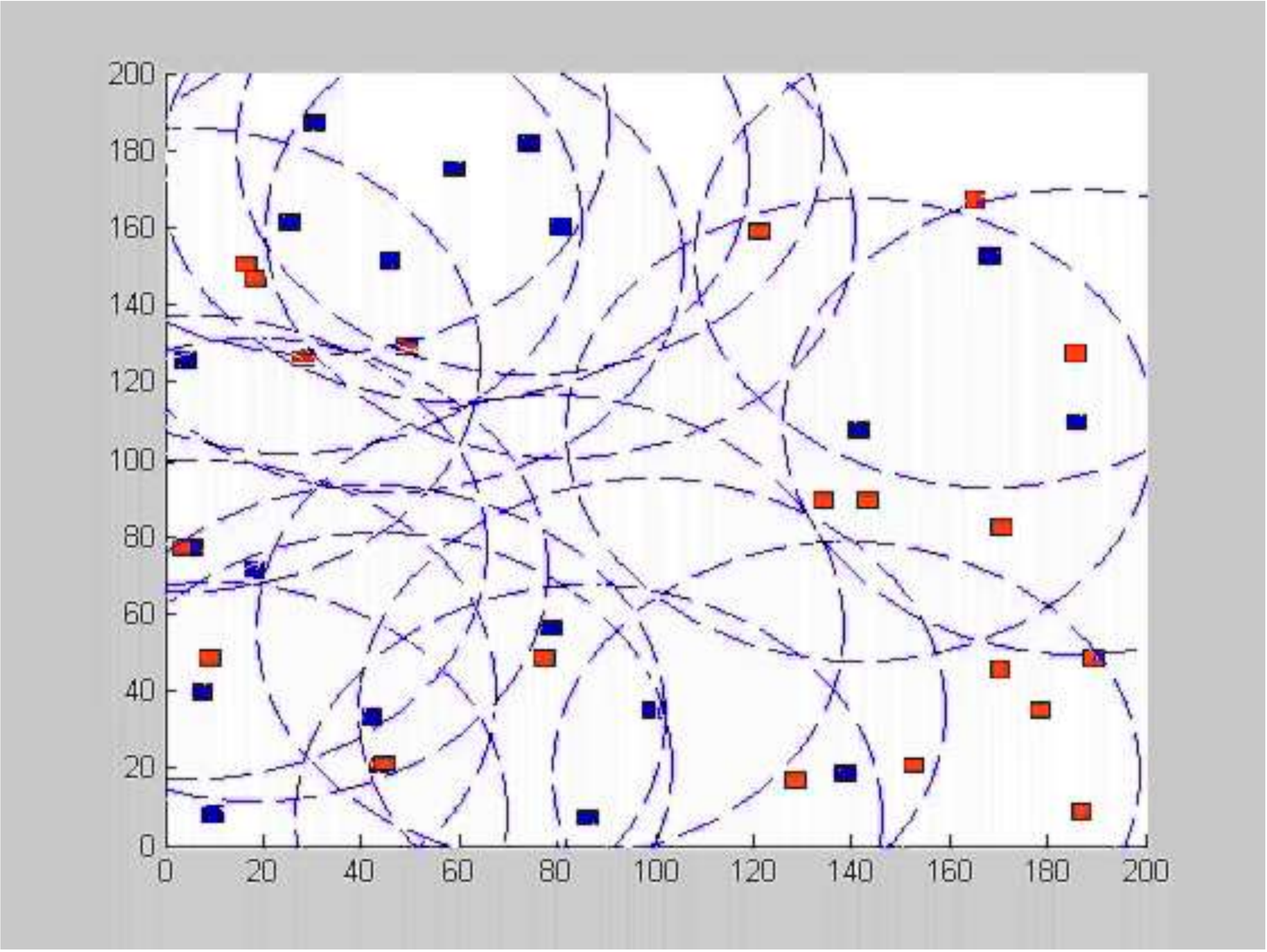
Near Optimum Performance

- **Example:**

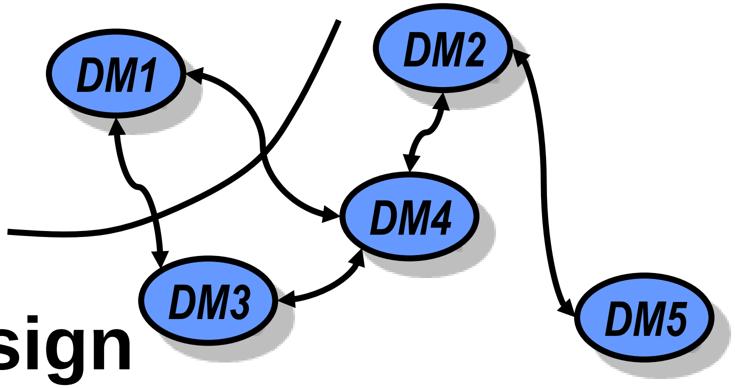
  200 uniform weapons negotiate 200 non-uniform targets



Greedy sequential implementation



Recap and Future Work



- Cooperative systems design
  - Agent utility design
  - Negotiation mechanism design

- Multiagent systems ripe for cross disciplinary efforts
  - Mission Planning with Autonomous Vehicles
  - Sensor Networks
  - Decentralized Inventory Control for Supply Chain Management

Design with Look-ahead …

(ongoing work)

Optimization
in Dynamic & Stochastic Environments

- Repeated decisions to optimize a long-term global utility

$$J_g = E \sum_{t=0}^{T} \alpha^t U_g(a(t), x(t))$$

- State x (t) changes stochastically as a function of a (t)

- State x (t) observed (partially) before choosing a (t),

$\Longrightarrow$ optimize over CL strategies $a(t) = \mu(x(t))$

THANK YOU !