GAME THEORY AND FOREIGN POLICY DECISION MAKING

David H. Wolpert, NASA Ames Research Center and Stanford

(david.h.wolpert@nasa.gov, ti.arc.nasa.gov/people/dhw/)

Stefan Bieniawski, Boeing

Juergen Jost, Max Planck Institute

Ritchie Lee, Carnegie Mellon University

Michael Harre, University of Sydney

TALK THEME

Combine

- ***Machine learning***
- ***Control theory***
- ***Behavioral game theory***

to control systems of *real-world* (!) interacting actors.

E.g., control interacting *people*, *firms*, *NGO's*, *states*, *non-state actors*

ROADMAP

1) Instantaneous control

No knowledge of system's environment

- Flutter suppression of a wing

↓

2) Instantaneous control

Detailed knowledge of system's environment

- Control pilots in near mid-air collisions

↓

3) Control of a full trajectory

No knowledge of system's environment

- Vary taxes to steer society to better equilibrium

FLUTTER SUPPRESSION IN AN AIRCRAFT WING

- **System**: ***Airplane wing with trailing edge microflaps***
- **Agents**: ***The microflaps, each running a separate Reinf. Learning (RL) controller (≅ a human)***

---

1) **Use game theory to predict system behavior for any set of {*agent reward functions*}**

2) Set reward functions to be learnable by the agents, and to result in desirable equilibrium behavior of the agents

FLUTTER SUPPRESSION IN AN AIRCRAFT WING

3) **Do not exploit how agent RL controllers work; only assume they work well.**

4) Have full power to set agent reward functions; sometimes unrealistic.

5) Did not account for dynamics; based on statics.

FLUTTER SUPPRESSION IN AN AIRCRAFT WING

ROADMAP

1) Instantaneous control

No knowledge of system's environment

- Flutter suppression of a wing

2) Instantaneous control

Detailed knowledge of system's environment

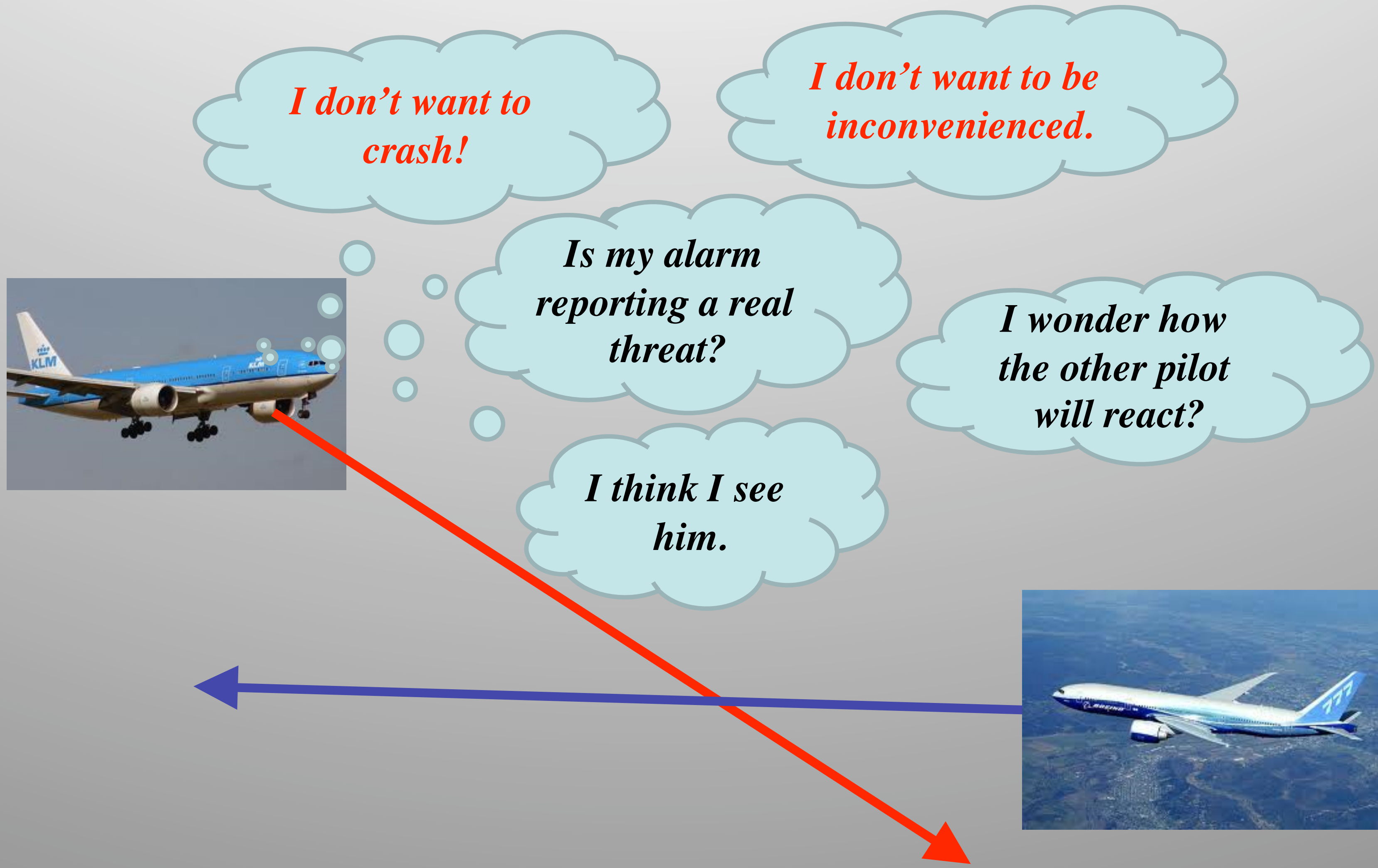
- Control pilots in near mid-air collisions

3) Control of a full trajectory

No knowledge of system's environment

- Vary taxes to steer society to better equilibrium

AVOIDING NEAR MID-AIR COLLISIONS



TCAS

Traffic Alert and Collision Avoidance System (TCAS)

- ***Monitors the airspace around an aircraft and provides Resolution Advisories (RA's) to pilots to avoid Mid-Air Collisions (MACs).***
- ***Based on a pilot model of perfect compliance with RA's***

A recent study of the Boston area…

<u>Traffic Alert and Collision Avoidance System (TCAS) compliance statistics of pilots:</u>

- ***13% Fully Compliant (pilots met assumptions about vertical speed, and promptness)***
- ***64% Partial Compliance (pilots moved in the proper direction, but not as promptly or aggressively as RA stipulated)***
- ***23% Non-compliance (pilots moved in the <u>OPPOSITE</u> direction to what RA stipulated)***

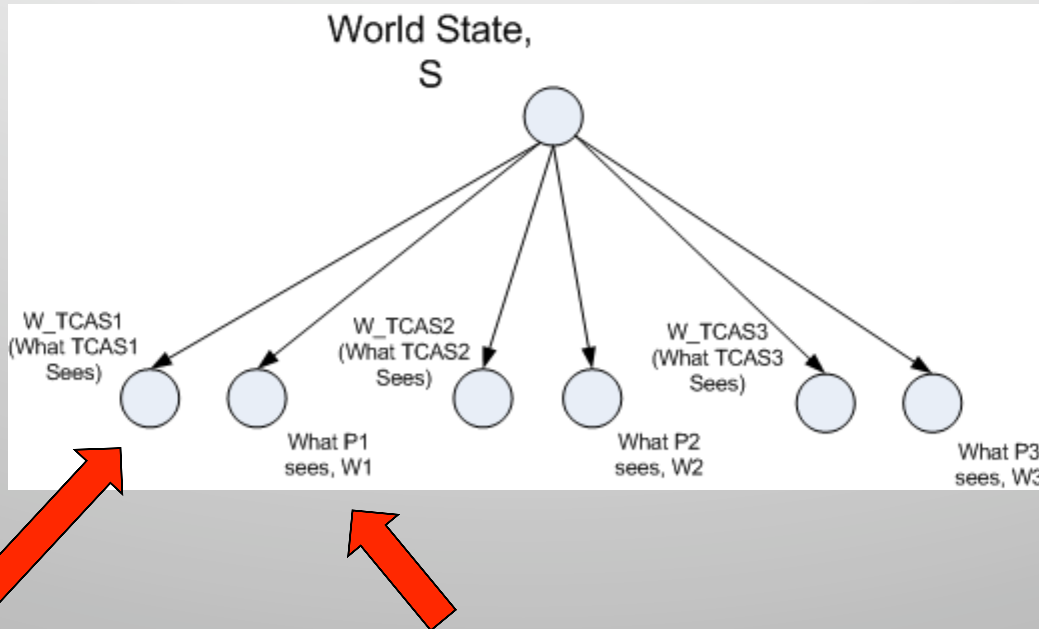
Source: Kuchar and Drumm

Bayes Net Model of an Encounter -1

Dynamic states of all aircraft

Bayes Net of an Encounter - 2

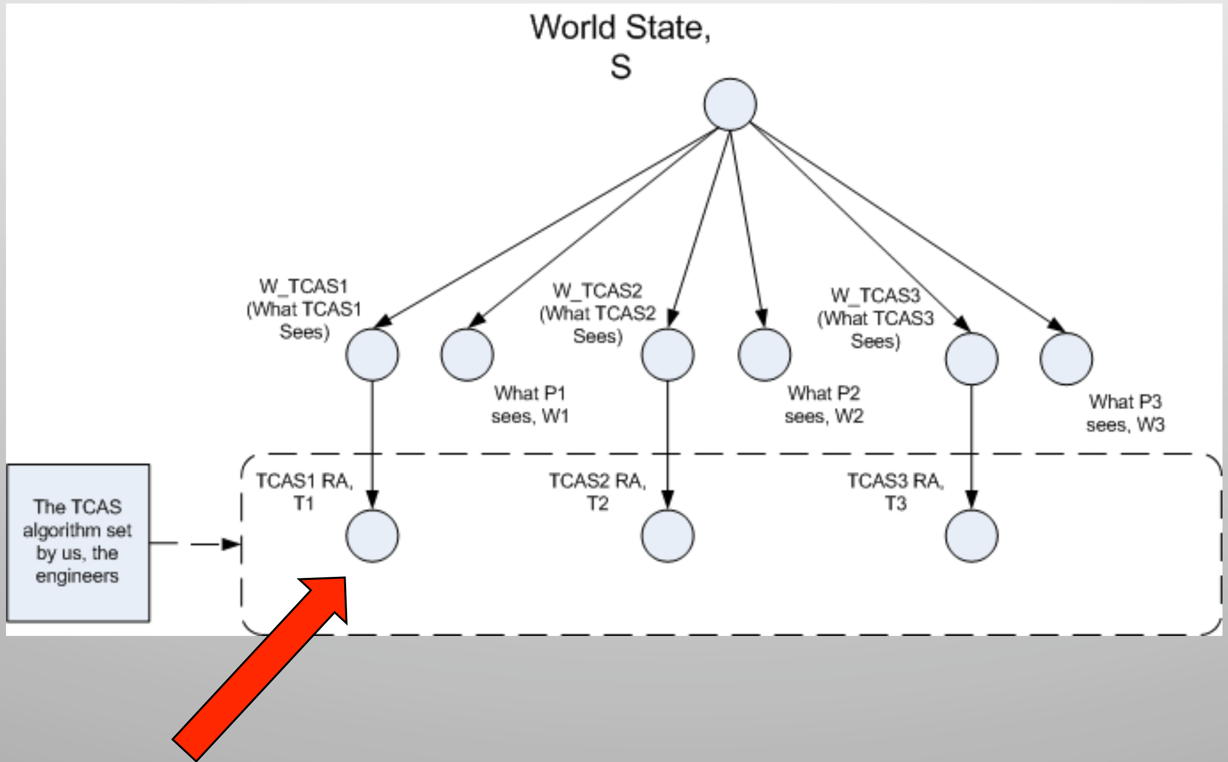
For 3 players:



W_TCAS1 is Player 1's TCAS computer's observation of the world

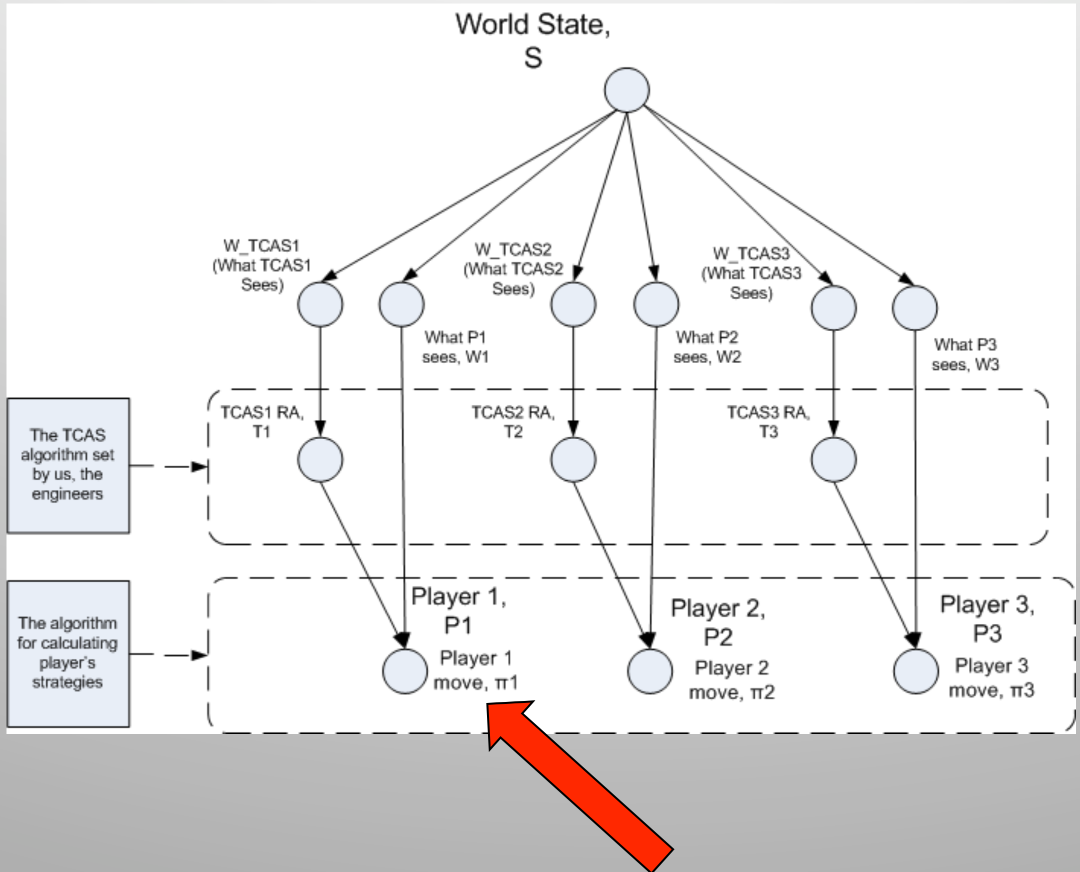
W1 is Player 1's observation of the world

Bayes Net of an Encounter - 3



T1 is the RA issued by Player 1's TCAS

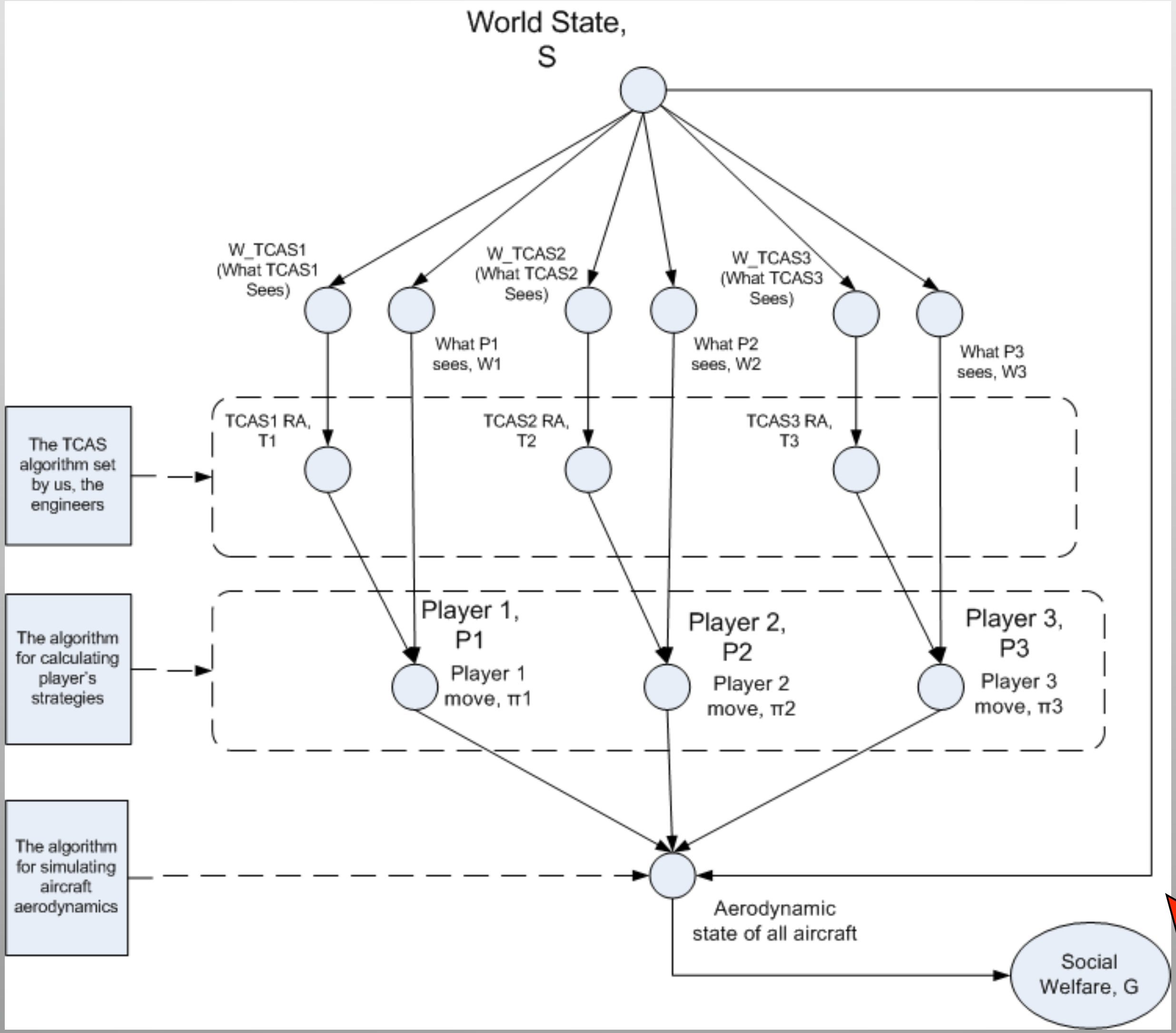
Bayes Net of an Encounter - 4



3 Terms in the Player's Utility Function

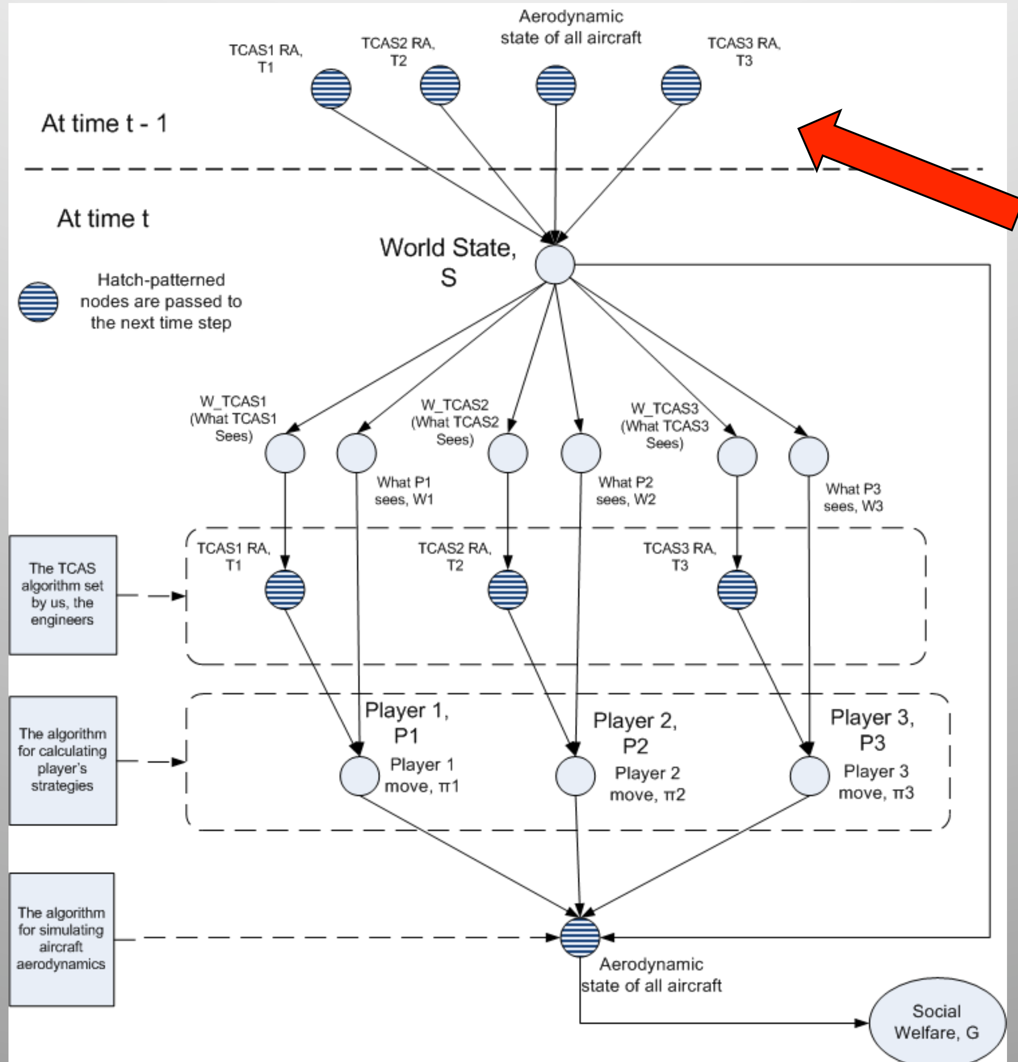
π1, Player 1's "move," is his desired vertical speed

Bayes Net of an Encounter - 5



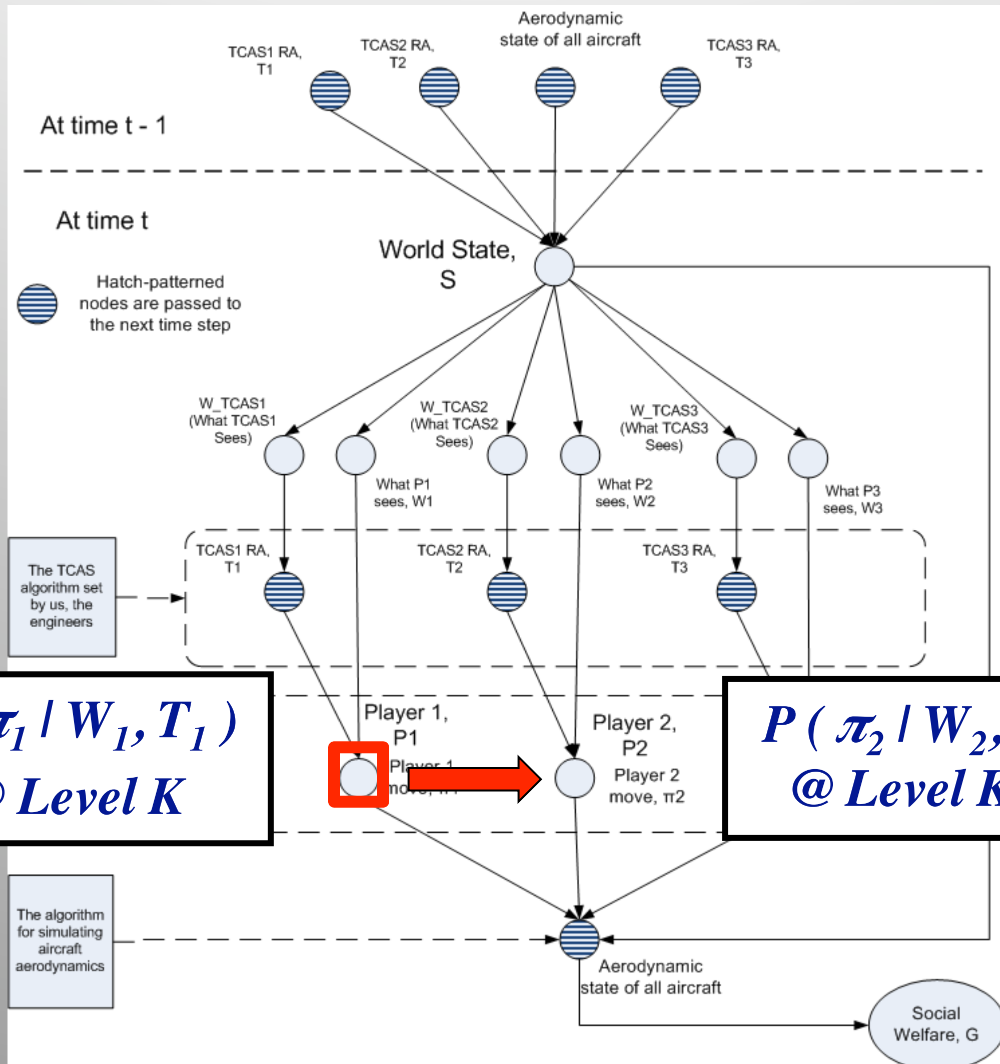
G is social welfare, taken as the minimum distance between aircraft in an encounter.

Bayes Net Time Extension



TCAS coordination and aircraft dynamic states are evolved and passed to the next time step

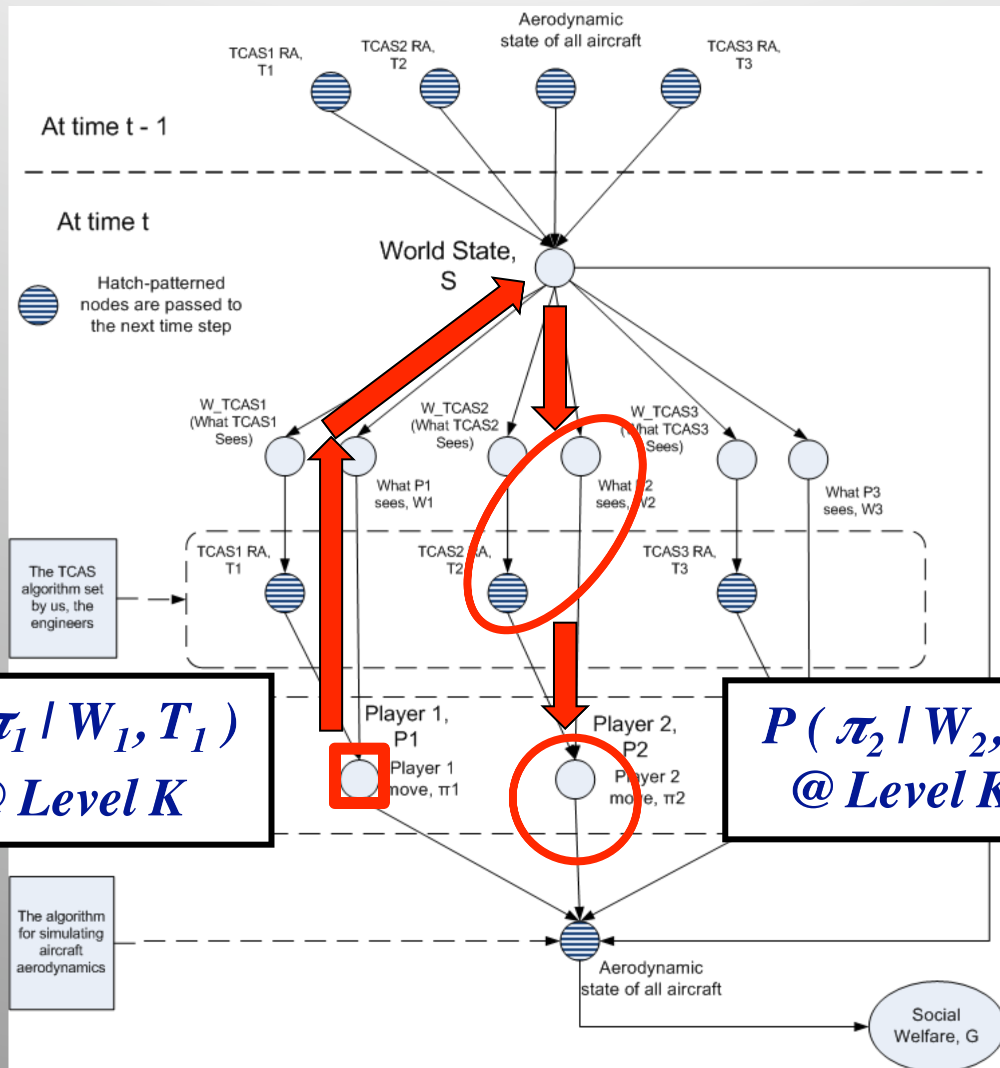
Bayesian Inversion



$P(\pi_1 \mid W_1, T_1)$
@ Level K

$P(\pi_2 \mid W_2, T_2)$
@ Level K - 1

Bayesian Inversion - 2

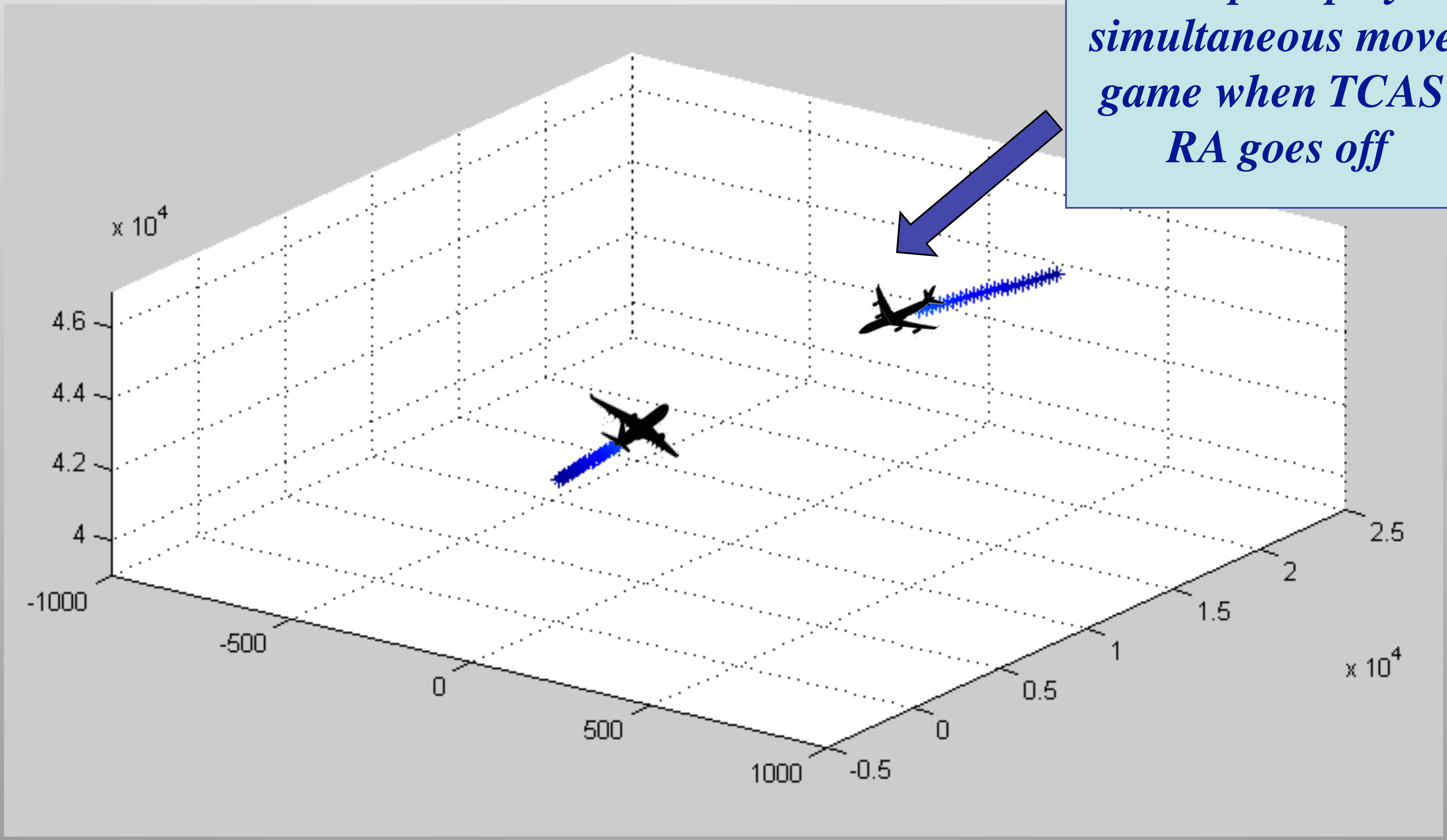


$P(\pi_1 \mid W_1, T_1)$
@ Level K

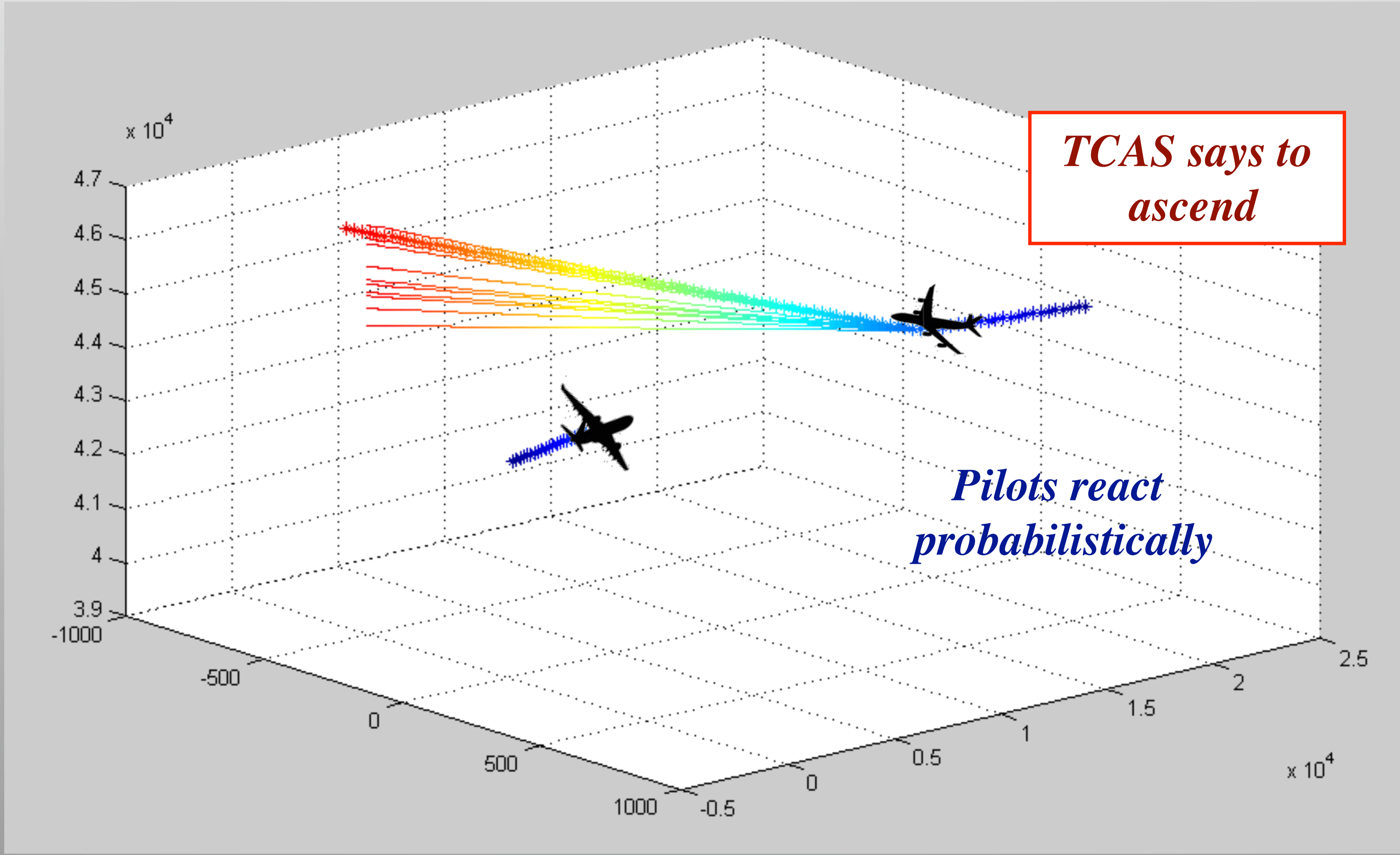
$P(\pi_2 \mid W_2, T_2)$
@ Level K - 1

Example Encounter

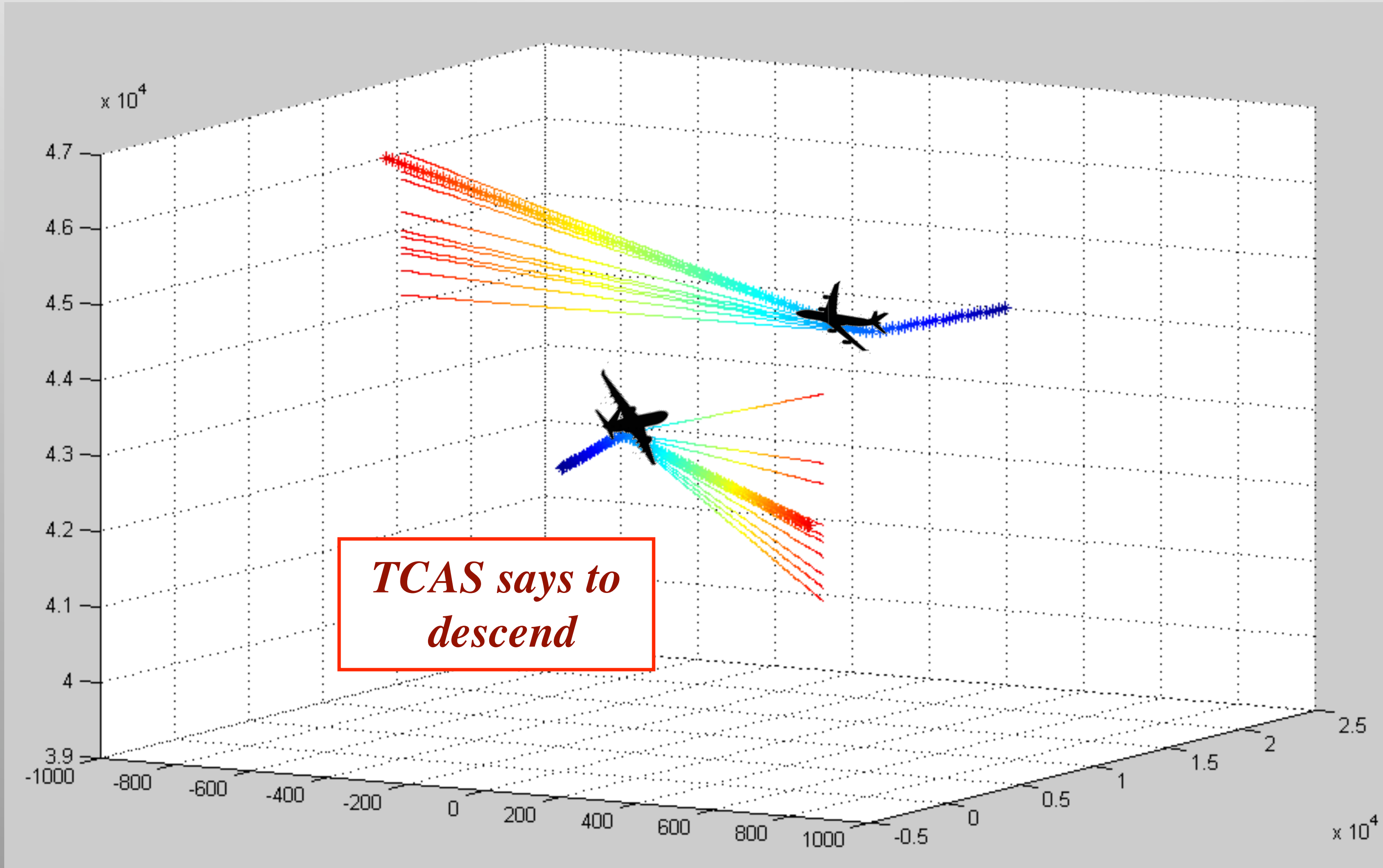
Each pilot plays simultaneous move game when TCAS RA goes off



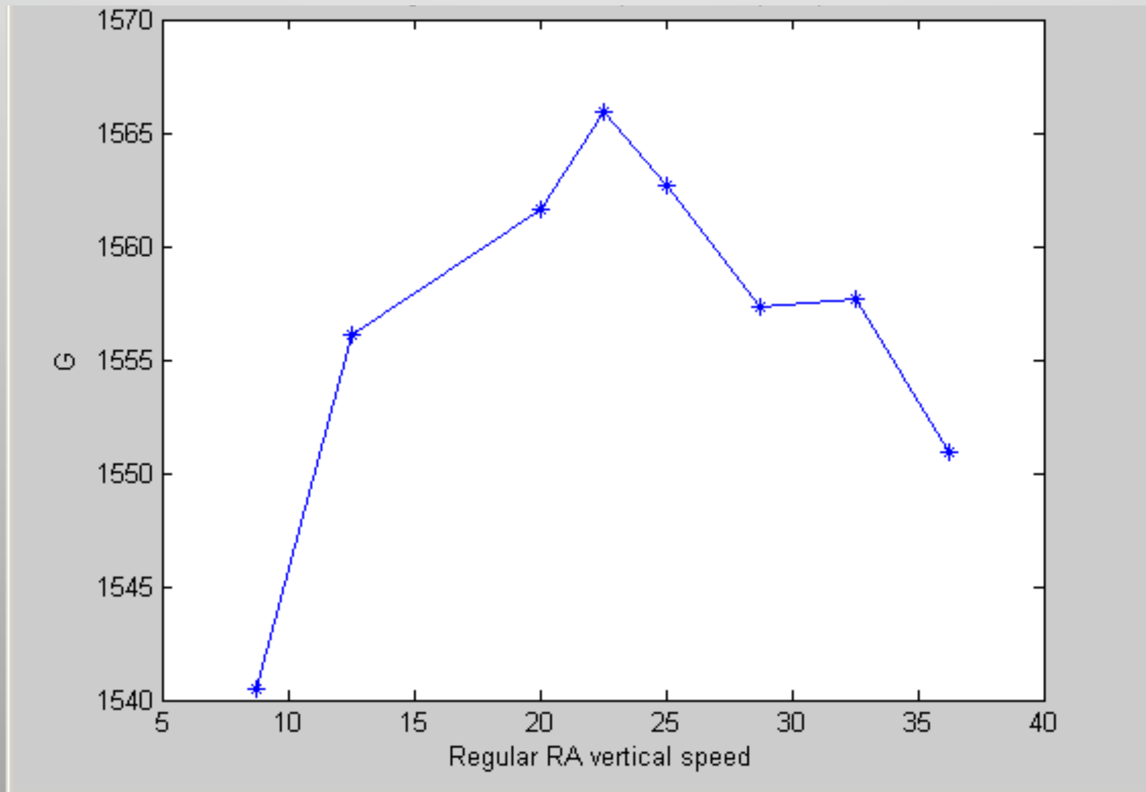
Example Encounter - 2



Example Encounter - 3



Social Welfare as a function of a TCAS Parameter (Regular RA Vertical Speed)



ROADMAP

1) Instantaneous control

No knowledge of system's environment

- Flutter suppression of a wing

2) Instantaneous control

Detailed knowledge of system's environment

- Control pilots in near mid-air collisions

3) Control of a full trajectory

No knowledge of system's environment

- Vary taxes to steer society to better equilibrium

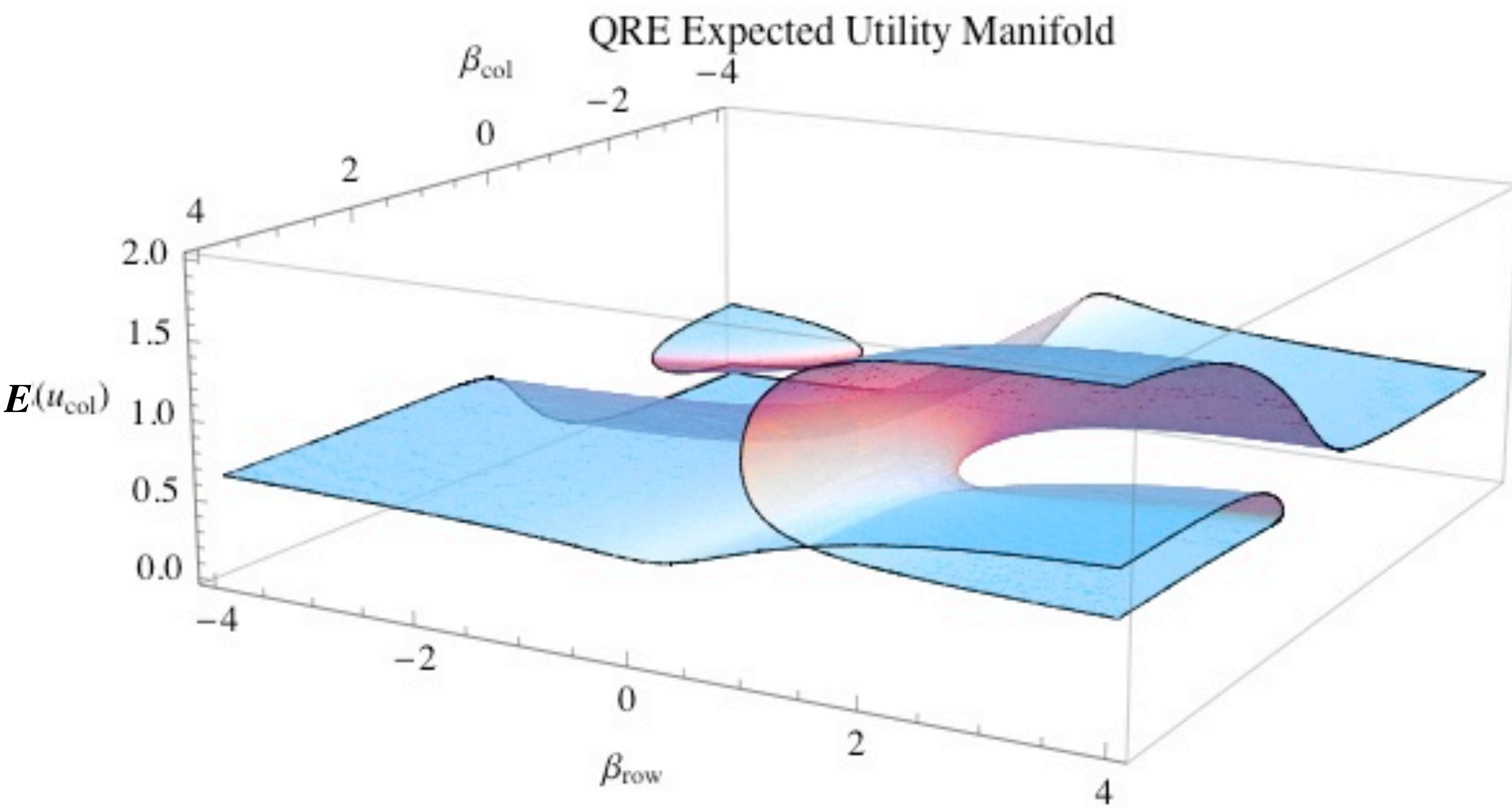
CONTROLLING SOCIETIES

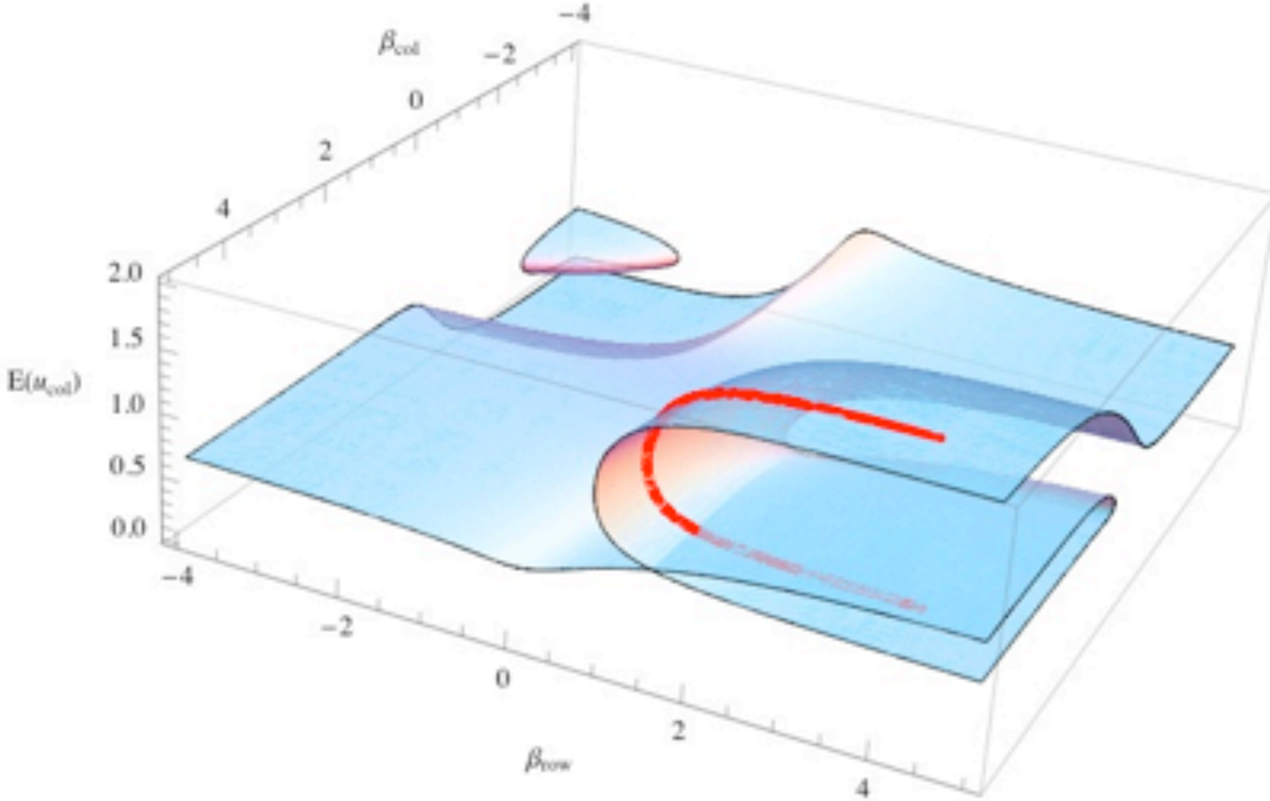
Example: A 2-player game with 2 moves per player, and the following pair of utility functions:

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

- **Players either artificial or natural (e.g., humans, firms)**
- **How do expected utilities in the associated QRE depend on the parameter vector (β_{row}, β_{col})?**

CONTROLLING SOCIETIES





- ***For a QRE, $q_i(x_i) \propto exp[\beta_i E(u_i \mid x_i)]$***

- ***So multiplying β_i by a factor a_i is equivalent to leaving β_i unchanged but multiplying u_i by a_i, i.e., to setting a tax rate of $1 - a_i$ on agent i***

- ***So same surface gives dependence of expected utilities on (individualized) agent tax rates, for fixed rationalities.***

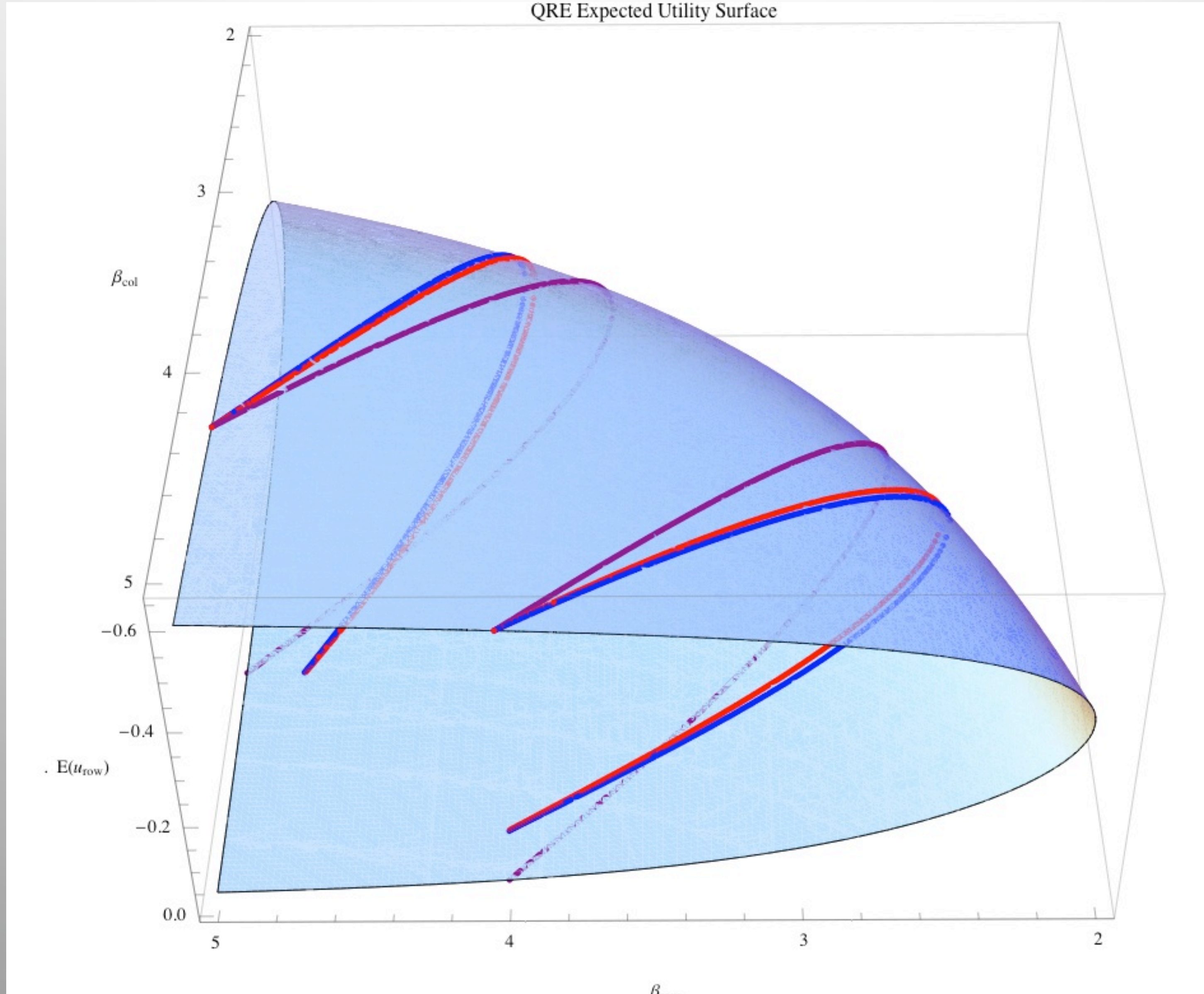
CONTROLLING SOCIETIES

Do there exist paths in (β_{row}, β_{col}) along which:

1) Neither player is ever more rational (has higher tax rate) than they are at the starting point;

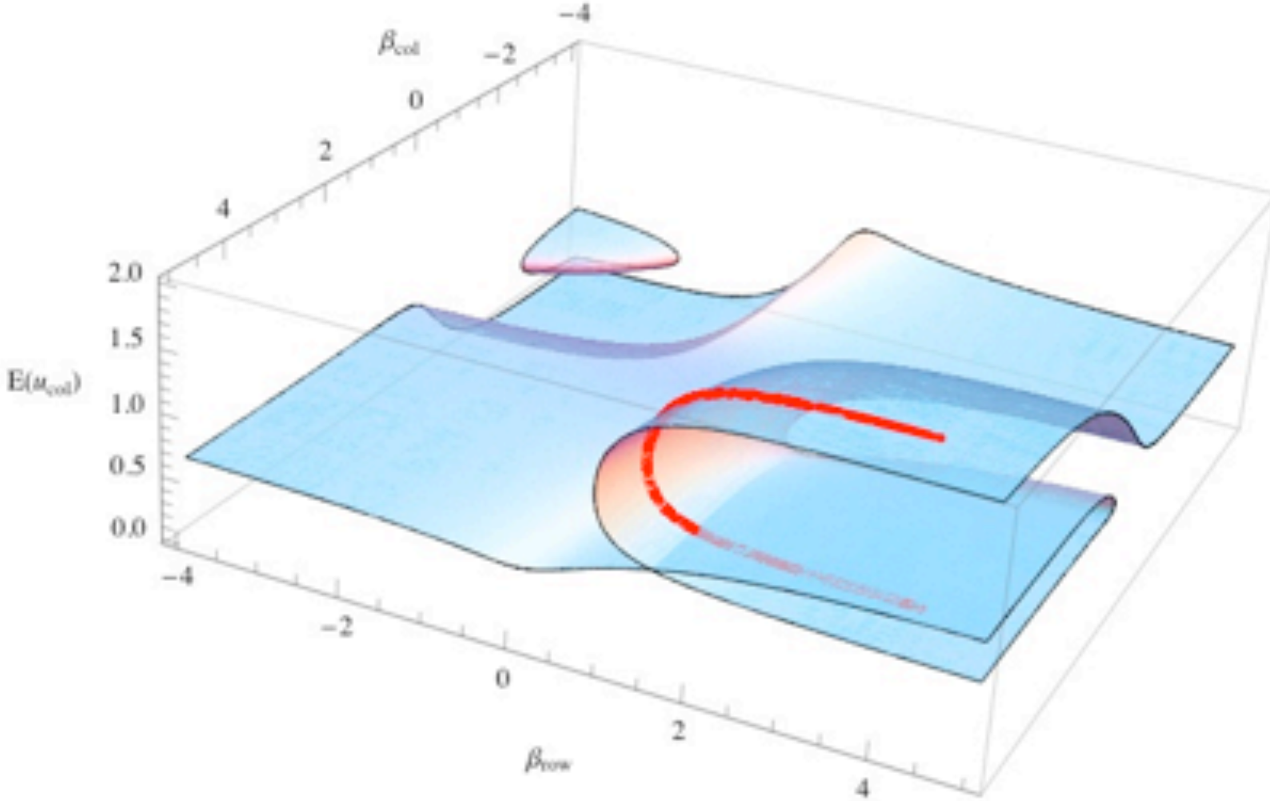
2) Both players increase expected utility at each step;

3) When possible, each player makes small rather than large strategy changes, and improving rather than hurting ones?



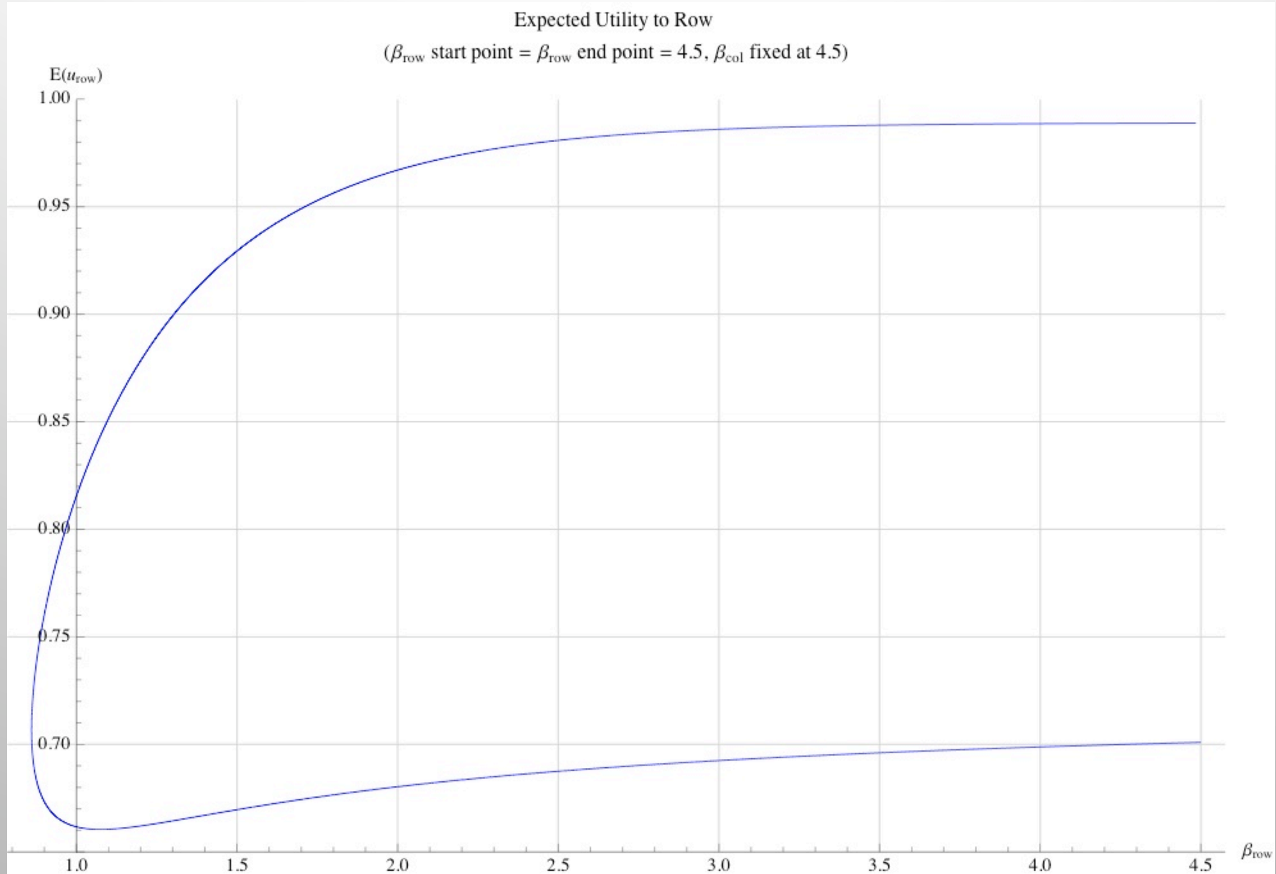
QRE Expected Utility Surface
β_{col}
2
3
4
5
−0.6
−0.4
−0.2
0.0
$E(u_{row})$
5
4
3
2
β_{row}

Yes.



- *At beginning of path, society needs to subsidize Row player a small amount to reduce β_{row}.*

- *When (β_{row}, β_{col}) returns to starting value, both players have gained expected utility.*



- ***At beginning of path, society needs to subsidize Row player a small amount to reduce β_{row}.***

- ***When (β_{row}, β_{col}) returns to starting value, even Row player has gained expected utility.***

CURRENT PROJECTS

Learn Probabilities and Utility Functions From Data

1) **Data types**:

- ***Real-world data (almost non-existent)***
- ***HITL data (very little – "high fidelity")***
- ***Crowd-sourced flight simulator data (huge amount, but "low fidelity")***

2) **Multi-fidelity approaches**:
Using statistical relation between different data types, exploit low-fidelity data to extend hi-fi data.

CURRENT PROJECTS

Automated analysis of Man-Machine Systems

1) Proposer of a man-machine system submits model of it

2) Use the model to estimate Pr(social welfare) for system

3) Use that to recommend tests of system, e.g.,
- to reduce error bars concerning battle outcomes, i.e., quantitative design of war games
- to verify (e.g., air traffic control system) for further development

CONCLUSION

Combine

- ***Machine learning***
- ***Control theory***
- ***Behavioral game theory***

to control systems of *real-world* (!) interacting actors.

E.g., control interacting *people*, *firms*, *NGO's*, *states*, *non-state actors*