Workshop on
Adversarial Decision Making

Adversarial Risk Analysis for Counterterrorism Modeling

Jesus Rios

IBM research

joint work with David Rios Insua

DIMACS, September 2010

Outline

- Motivation
- ARA framework:
  Predicting actions from intelligent others
- (Basic) counterterrorism models
  - Sequential Defend-Attack model
  - Simultaneous Defend-Attack model
  - Defend-Attack-Defend model
  - Sequential Defend-Attack model with Defender's private info.
- Discussion

Motivation

- Biological Threat Risk Analysis for DHS (Battelle, 2006)
  - Based on Probability Event Trees (PET)
    - Government & Terrorists' decisions treated as random events

- Methodological improvements study (NRC committee)
  - PET appropriate for risk assessment of
    - Random failure in engineering systems

    but not for adversarial risk assessment
    - Terrorists are intelligent adversaries trying to achieve their own objectives
    - Their decisions (if rational) can be somehow anticipated
  - PET cannot be used for a full risk management analysis
    - Government is a decision maker not a random variable

Methodological improvement recommendations

- Distinction between risk from
  - Nature/Chance vs.
  - Actions of intelligent adversaries

- Need of models to predict Terrorists' behavior
  - Red team role playing (simulations of adversaries thinking)
  - Attack-preference models
    - Examine decision from Attacker viewpoint (T as DM)
  - Decision analytic approaches
    - Transform the PET in a decision tree (G as DM)
      - How to elicit probs on terrorist decisions??
      - Sensitivity analysis on (problematic) probabilities
        - » Von Winterfeldt and O'Sullivan (2006)
  - Game theoretic approaches
    - Transform the PET in a game tree (G & T as DM)

Adversarial risk problems

- Two (or more) intelligent opponents
  - Defender invests in a portfolio of defense options
  - Terrorists invest effort and distribute resources among different types of attack
- Uncertain outcomes
  - arising both from randomness and our lack of knowledge
- Advise the Defender to efficiently spend resources
  - To reduce/eliminate the risks from malicious (or self-interested) actions of intelligent adversaries

Tools for analysis

- Chance and uncertainty analysis
  - Statistical risk analysis
    - Terrorists’ actions as a random variables
- Decision making paradigms
  - Game theory (multiple DMs)
    - Terrorists’ actions as a decision variables
  - Decision Analysis (unitary DM)
    - Terrorists’ actions as a random variables
- Graphical representations
  - Game and decision trees
  - Multi-agent Influence Diagrams

Critiques to the Game Theoretic approach

- Unrealistic assumptions
  - **Full and common knowledge assumption**
    - e.g. Attacker's objectives are known
  - Common prior assumption for games with private information
- Symmetric predictive and descriptive approach
  - What if multiple equilibria
  - Passive understanding
- Equilibria does not provide partisan advise
- Impossibility to accommodate all kind of information that may be available (intelligence about what the attacker might do)

Decision analytic approaches

- One-sided prescriptive support
  - Use a prescriptive model (SEU) for supporting the Defender
  - Treat the Attacker's decision as uncertainties
  - Help the Defender to assess probabilities of Attacker's decisions

- The 'real' bayesian approach to games (Kadane & Larkey 1982)
  - Weaken common (prior) knowledge assumption

- Asymmetric prescriptive/descriptive approach (Raiffa 2002)
  - Prescriptive advice to one party conditional on a (probalistic) description of how others will behave

- Adversarial Risk Analysis
  - Develop methods for the analysis of the adversaries' thinking to anticipate their actions.
    - We assume the Attacker is a *expected utility maximizer*
    - But other (*descriptive)* models may be possible

Predicting actions from intelligent others

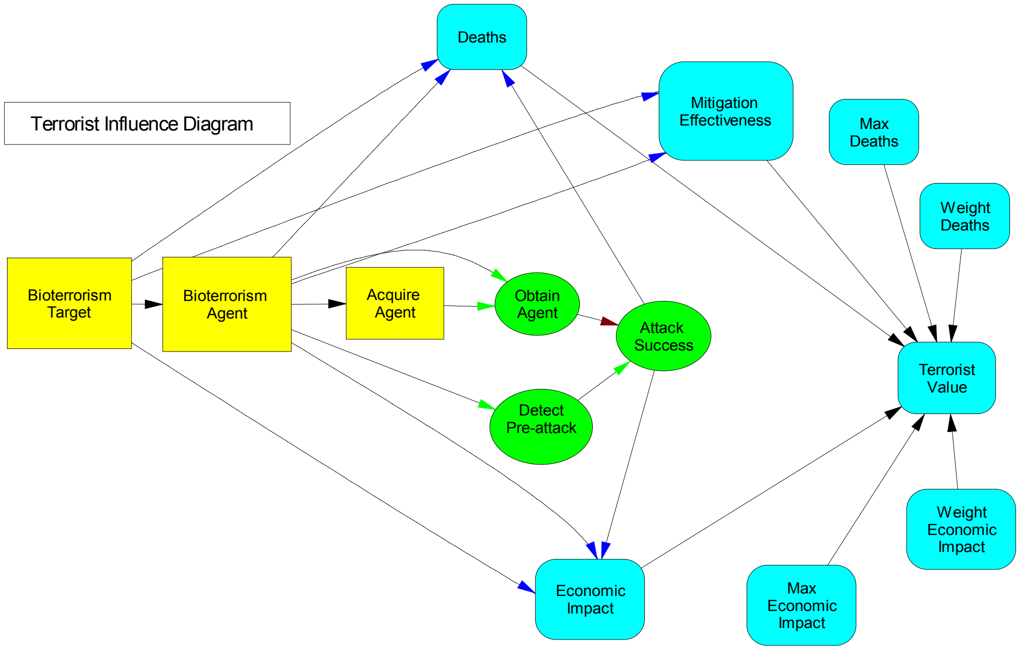
- Decision analytic approach
  - Prob over the actions of intelligent others
  - Compute defence of maximum expected utility

- How to assess a probability distribution over the actions (attacks) of an intelligent adversary??

- (Probabilistic) modeling of terrorist’s actions
  - Attack-preference models
    - Examine decision from Attacker viewpoint

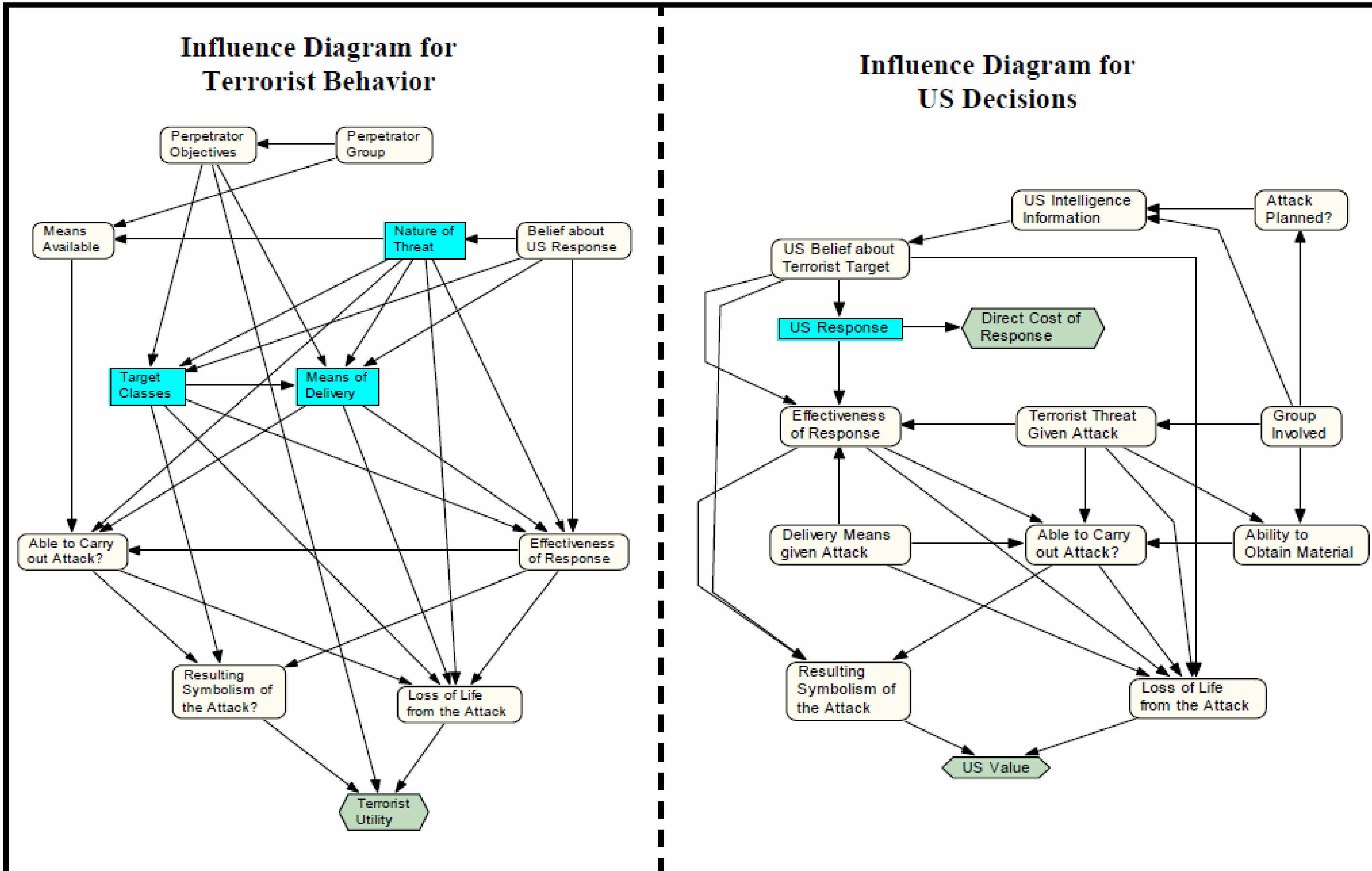
Parnell (2007)

- Elicit Terrorist's probs and utilities from our viewpoint
  - Point estimates
- Solve Terrorist's decision problem
  - Finding Terrorist's action that gives him max. expected utility
- Assuming we know the Terrorist's true probs and utilities
  - We can anticipate with certitude what the terrorist will do



Paté-Cornell & Guikema (2002)

Attacker *Defender*



Paté-Cornell & Guikema (2002)

- Assessing probabilities of terrorist's actions
  - From the Defender viewpoint
    - Model the Attacker's decision problem
    - Estimate Attacker's probs and utilities
    - Calculate expected utilities of attacker's actions
  - Prob of attacker's actions proportional to their perceived expected utilities

- Feed with these probs the uncertainty nodes with Attacker's decisions in the Defender's influence diagram
  - Choose defense of maximum expected utility

- Shortcoming
  - If the (idealized) adversary is an *expected utility maximizer* he would certainly choose the attack of max expected utility
  - a choice that could be divined by the analyst, if the analyst knows the adversary's true utilities and risk analysis

How to assess probabilities over the actions of an intelligent adversary??

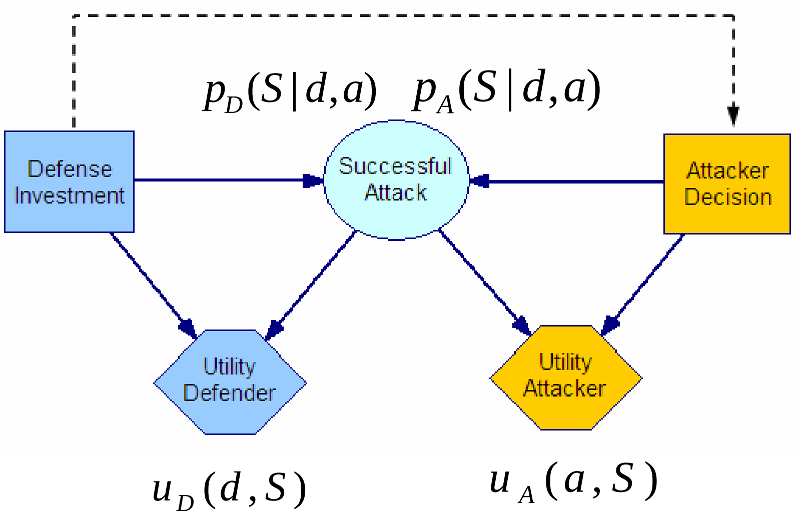
- Raiffa (2002): Asymmetric prescriptive/descriptive approach
  - Lab role simulation experiments
  - Assess probability distribution from experimental data
- Our proposal: Rios Insua, Rios & Banks (2009)
  - Assessment based on an analysis of the adversary rational behavior
    - Assuming the Attacker is a SEU maximizer
      - Model his decision problem
      - Assess his probabilities and utilities
      - Find his action of maximum expected utility
  - Uncertainty in the Attacker's decision stems from
    - *our* uncertainty about his probabilities and utilities
  - Sources of information
    - Available past statistical data of Attacker's decision behavior
    - Expert knowledge / Intelligence
    - Non-informative (or reference) distributions

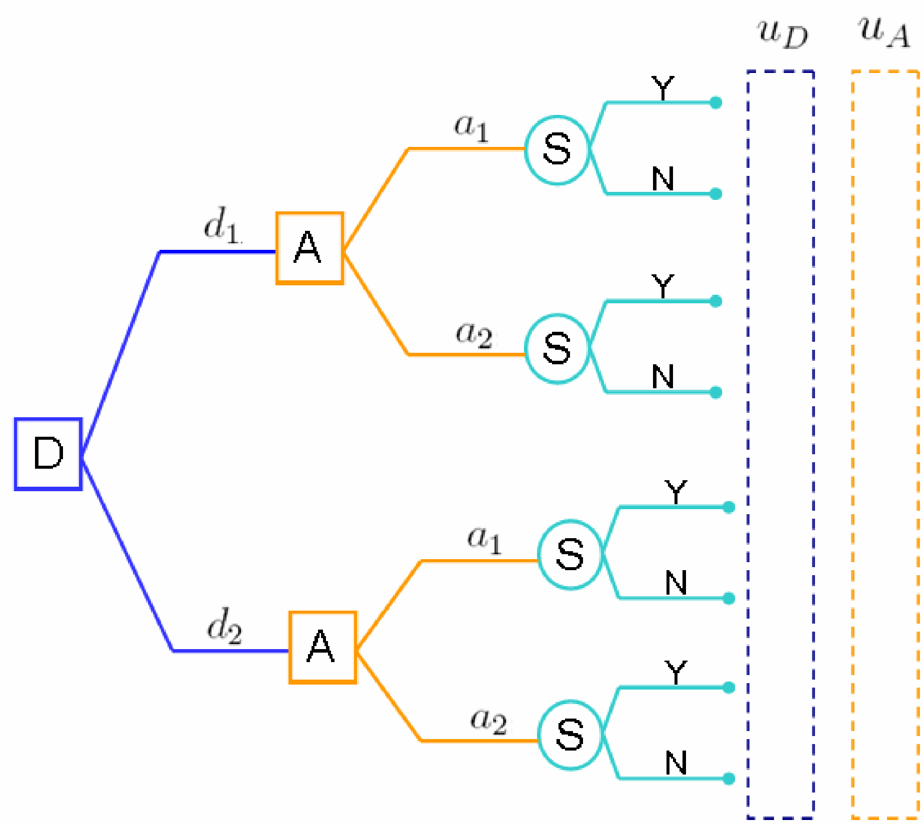
Counterterrorism modeling

- Basic models
- Standard Game Theory vs. Bayesian Decision Analysis
- Supporting the Defender against an Attacker
- How to assess Attacker's decisions (probability of Attacker's actions)
  - No infinity regress
    - sequential Defender-Attacker model
  - Infinity regress
    - simultaneous Defender-Attacker model

Sequential Defend-Attack model

- Two intelligent players
  - Defender and Attacker
- Sequential moves
  - First Defender, afterwards Attacker knowing Defender's decision





Standard Game Theoretic Analysis

Expected utilities at node S

$$\psi_D(d,a) = p_D(S=0|d,a)\ u_D(d,S=0)\ +\ p_D(S=1|d,a)\ u_D(d,S=1)$$

$$\psi_A(d,a) = p_A(S=0 \mid d,a)\ u_A(a,S=0)\ +\ p_A(S=1 \mid d,a)\ u_A(a,S=1)$$

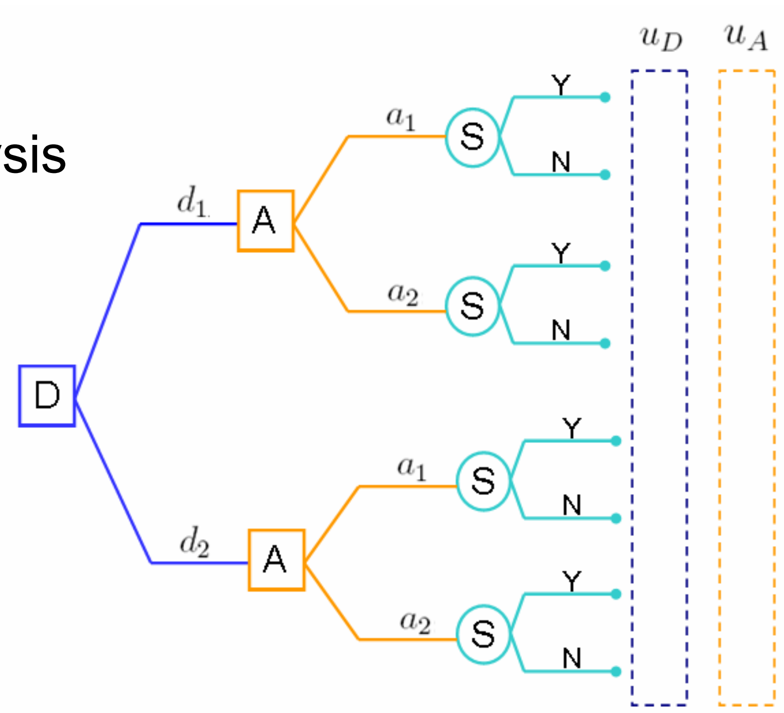
Best Attacker's decision at node A

$$a^*(d) = \operatorname{argmax}_{a \in \mathcal{A}} \psi_A(d,a)$$

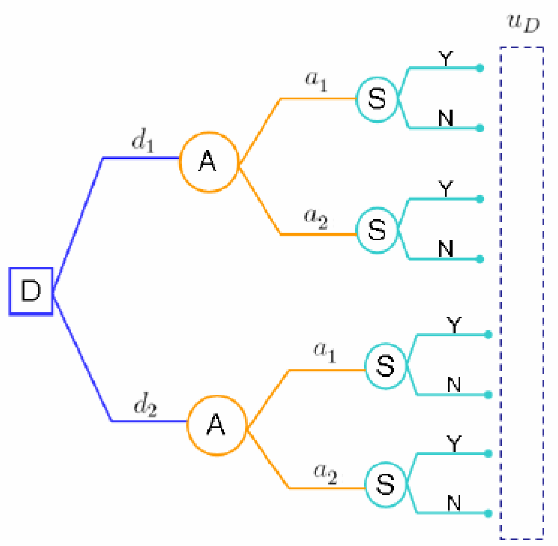
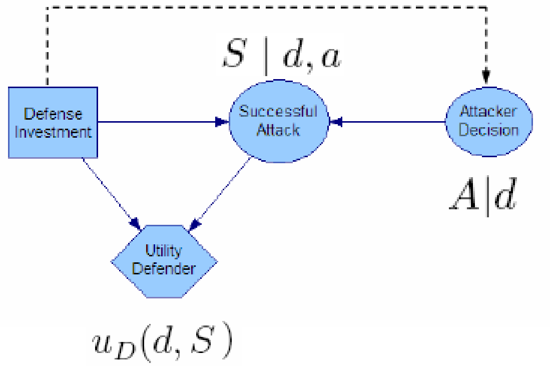
Assuming Defender knows Attacker's analysis
Defender's best decision at node D

$$d^* = \operatorname{argmax}_{d \in \mathcal{D}} \psi_D(d, a^*(d))$$

Solution: $(d^*, a^*(d^*))$



ARA: Supporting the Defender

Defender's problem	Defender's solution of maximum SEU
 	$\psi_D(d,a) = p_D(S=0 \mid d,a)\; u_D(d, S=0) \; + \; p_D(S=1 \mid d,a)\; u_D(d, S=1)$ $\psi_D(d) = \psi_D(d,a_1)\; p_D(A=a_1 \mid d) + \psi_D(d,a_2)\; p_D(A=a_2 \mid d)$ $d^* = \arg\max_{d \in X_D} \psi_D(d)$ Modeling input: $p_D(S \mid a,d)$ $p_D(A \mid d)$??

Example: Banks-Anderson (2006)

- Exploring how to defend US against a possible smallpox attack
  - Random costs (payoffs)

	No Attack	Minor Attack	Major Attack
Stockpile	C_{11}	C_{12}	C_{13}
Biosurveillance	C_{21}	C_{22}	C_{23}
First Responders	C_{31}	C_{32}	C_{33}
Mass Inoculation	C_{41}	C_{42}	C_{43}

  - Conditional probabilities of each kind of smallpox attack given terrorist knows what defence has been adopted

*This is
the problematic step
of the analysis*

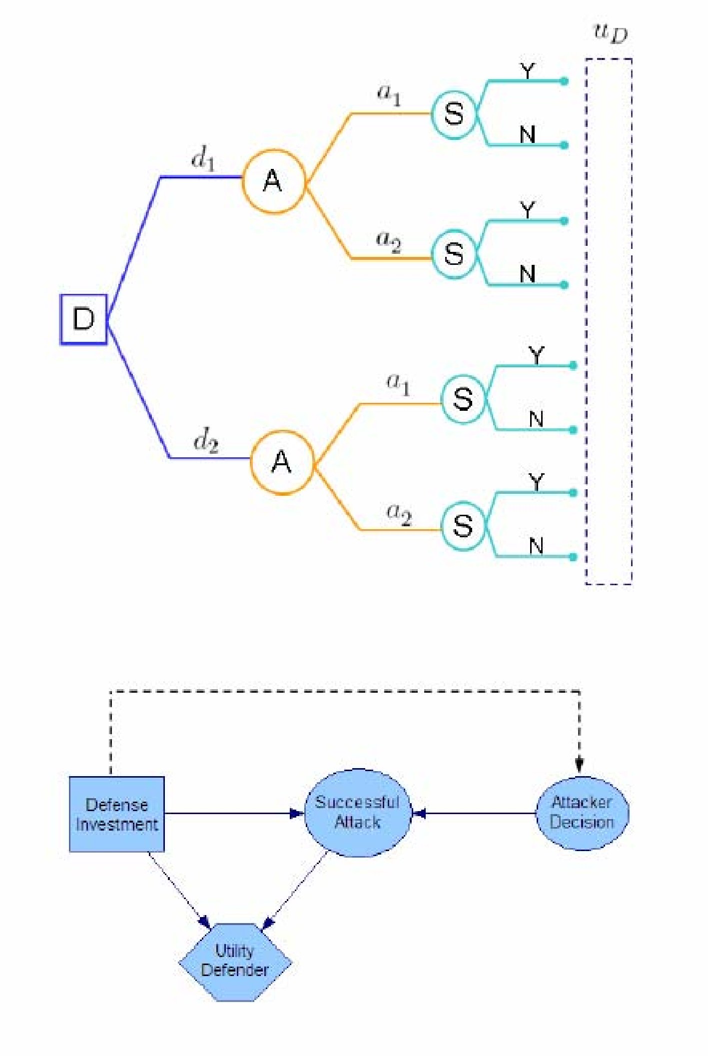
	No Attack	Minor Attack	Major Attack
Stockpile	.95	.040	.010
Biosurveillance	.96	.035	.005
First Responders	.96	.039	.001
Mass Inoculation	.99	.009	.001

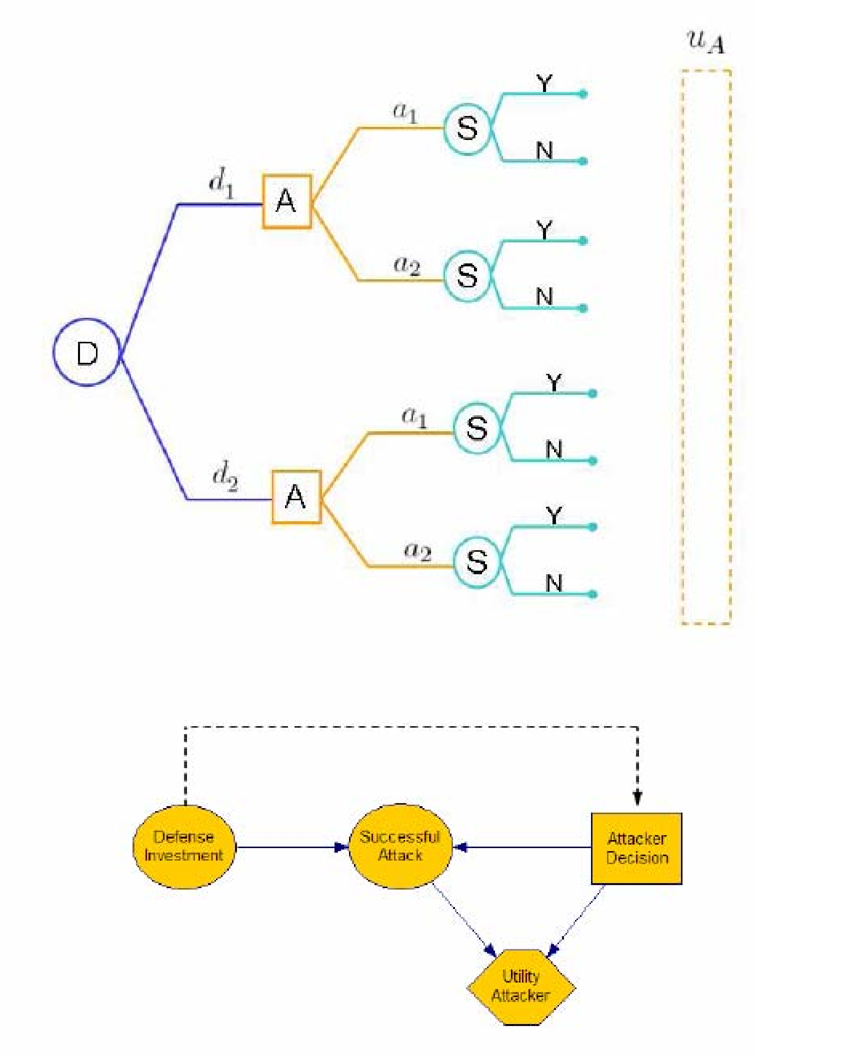
  - Compute expected cost of each defence strategy

- Solution: defence of minimum expected cost

Predicting Attacker's decision: $p_D(A \mid d)$

Defender problem	Defender's view of Attacker problem

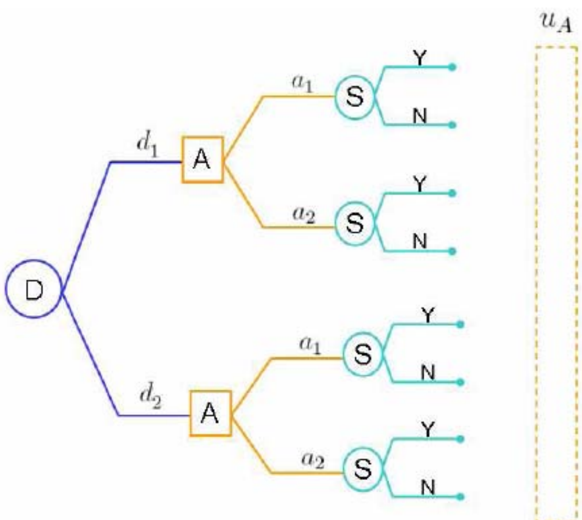


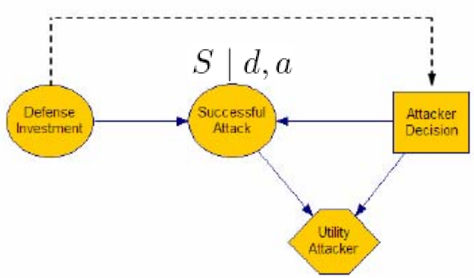


Solving the assessment problem

Defender's view of Attacker problem	Elicitation of $p_D(A\,|d)$

Defender's view of Attacker problem





Elicitation of $p_D(A\,|d)$

A is an EU maximizer

D's beliefs about $(u_A, p_A) \sim (P_A, U_A) = F$

$$\Psi_A(d,a) = P_A(S=0 \mid d,a)\, U_A(a, S=0) \;+\; P_A(S=1 \mid d,a)\, U_A(a, S=1)$$

$$p_D(A = a|d) = \mathbb{P}_F[a = \mathrm{argmax}_{x\in\mathcal{A}} \Psi_A(d,x)]$$

<u>MC simulation</u>

$$\{(p_A^i, u_A^i)\}_{i=1}^n \sim F \;\longrightarrow\; \psi_A^i \sim \Psi_A$$

$$a_i^*(d) = \mathrm{argmax}_{x\in\mathcal{A}}\, \psi_A^i(x,d)$$

$$p_D(A = a|d) \approx \#\{a = a_i^*(d)\}/n$$

Bayesian decision solution for the sequential Defend- Attack model

1. Assess (p_D, u_D) from the Defender

2. Assess $F = (P_A, U_A)$, describing the Defender's uncertainty about (p_A, u_A)

3. For each d, simulate to assess $p_D(A|d)$ as follows:

   (a) Generate $(p_A^i, u_A^i) \sim F$, $i = 1, \ldots, n$

   Solve $a_i^*(d) = \text{argmax}_{a \in \mathcal{A}}\ \psi_A^i(d, a)$

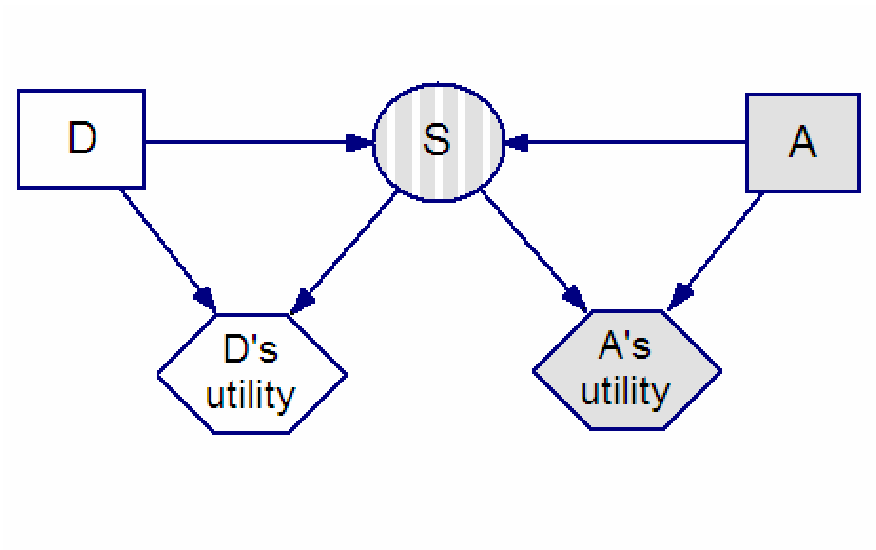
   (b) Approximate $\hat{p}_D(A = a|d) = \#\{a = a_i^*(d)\}/n$

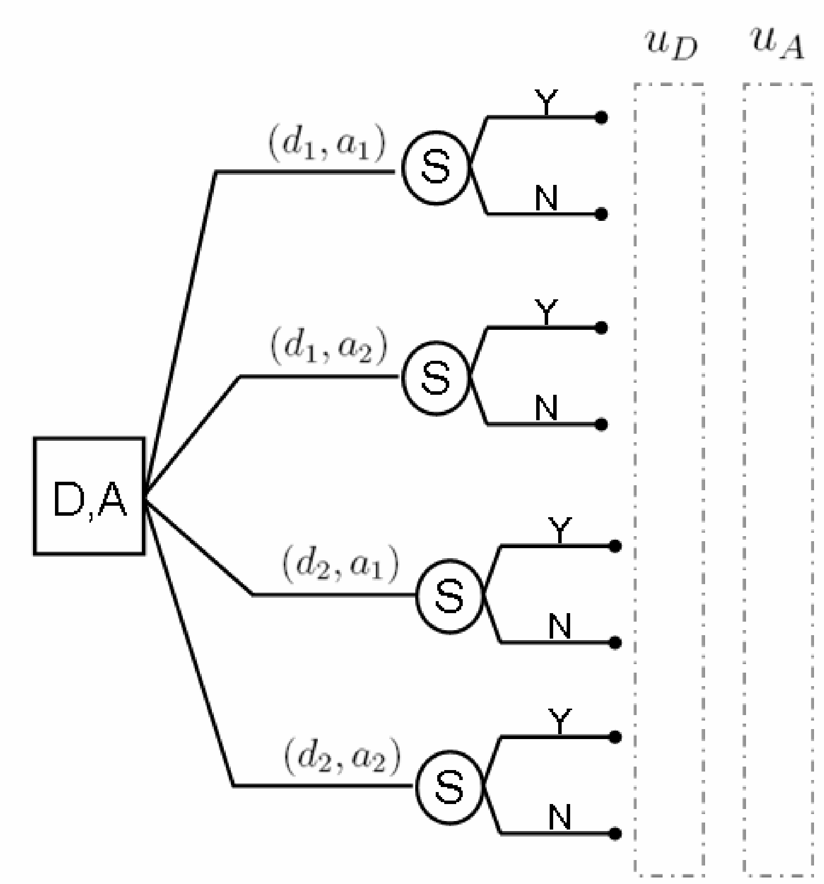
4. Solve the Defender's problem

$$d^* = \text{argmax}_{d \in \mathcal{D}}\ \psi_D(d, a_1)\ \hat{p}_D(A = a_1|d)\ +\ \psi_D(d, a_2)\ \hat{p}_D(A = a_2|d)$$

Simultaneous Defend-Attack model

- Decisions are taken without knowing each other's decisions





Game Theory Analysis

- Common knowledge
  - Each knows expected utility of every pair (d,a) for both of them
  - Nash equilibrium: (d*, a*) satisfying

$$\psi_D(d^*, a^*) \geq \psi_D(d, a^*) \;\; \forall\, d \in \mathcal{D}$$

$$\psi_A(d^*, a^*) \geq \psi_A(d^*, a) \;\; \forall\, a \in \mathcal{A}$$

- When some information is not common knowledge
  - Private information
    - Type of Defender and Attacker

$$\tau_D \in T_D \longrightarrow u_D(d, s, \tau_D) \quad p_D(S \mid d, a, \tau_D)$$

$$\tau_A \in T_A \longrightarrow u_A(d, s, \tau_D) \quad p_A(S \mid d, a, \tau_D)$$

  - Common prior over private information $\pi(\tau_D, \tau_A)$
  - Model the game as one of incomplete information

Bayes Nash Equilibrium

– Strategy functions
  - Defender $d : \tau_D \to d(\tau_D) \in \mathcal{D}$
  - Attacker $a : \tau_A \to a(\tau_A) \in \mathcal{A}$

– Expected utility of (d,a)
  - for Defender, given her type $\psi_D(d(\tau_D), a, \tau_D) =$

$$= \int \underbrace{\left[\sum_{s \in S} u_D(d(\tau_D), s, \tau_D)\, p_D(S = s \mid d(\tau_D), a(\tau_A), \tau_D) \right]}_{\psi_D(d(\tau_D), a(\tau_A), \tau_D)} \pi(\tau_A \mid \tau_D)\, \mathrm{d}\tau_A$$

  - Similarly for Attacker, given his type $\psi_A(d, a(\tau_A), \tau_A)$

– Bayes-Nash Equlibrium (d*, a*) satisfying

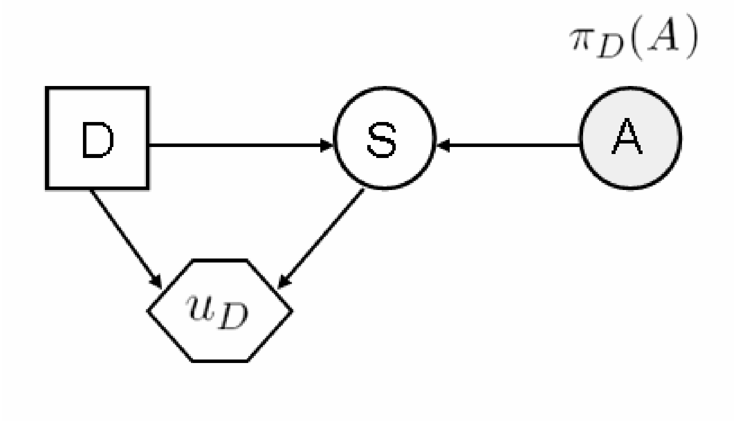
$$\psi_D(d^*(\tau_D), a^*, \tau_D) \geq \psi_D(d(\tau_D), a^*, \tau_D) \quad \forall\, d : \tau_D \to d(\tau_D)$$

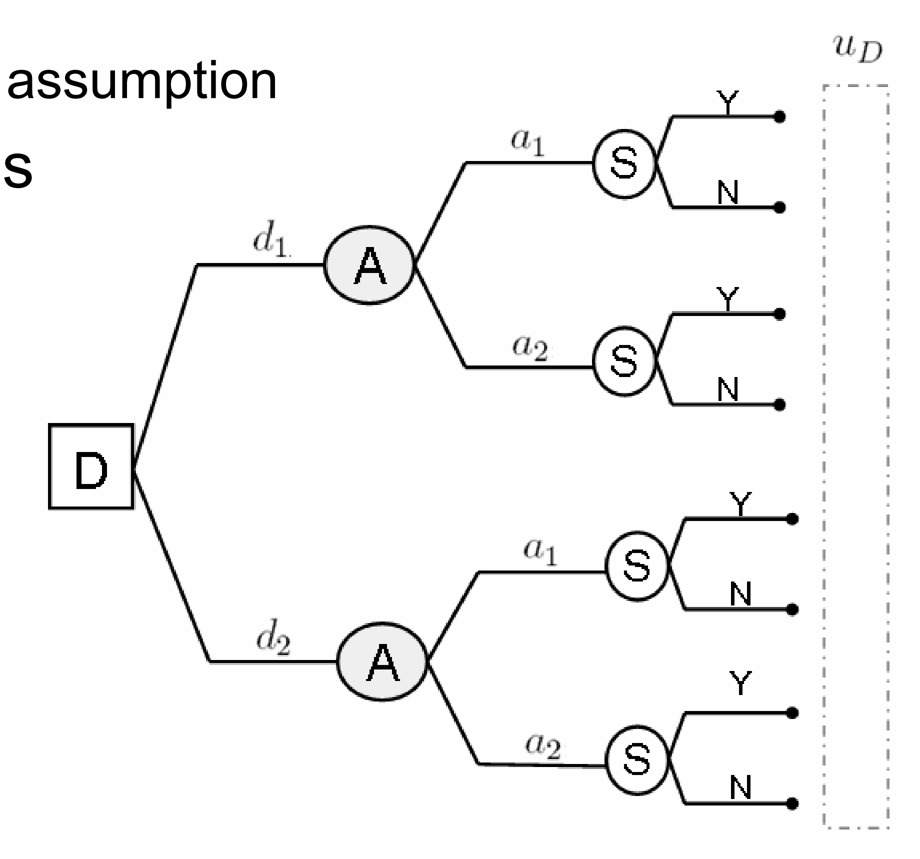
$$\psi_A(d^*, a^*(\tau_A), \tau_A) \geq \psi_A(d^*, a(\tau_A), \tau_A) \quad \forall\, a : \tau_A \to a(\tau_A)$$

ARA: Supporting the Defender

Weaken common (prior) knowledge assumption

- Defender's decision analysis



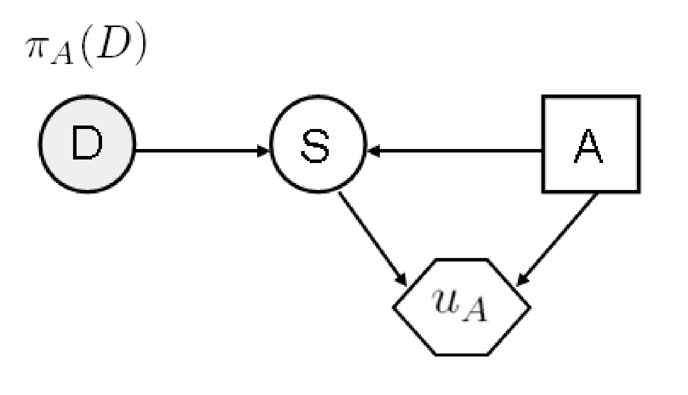


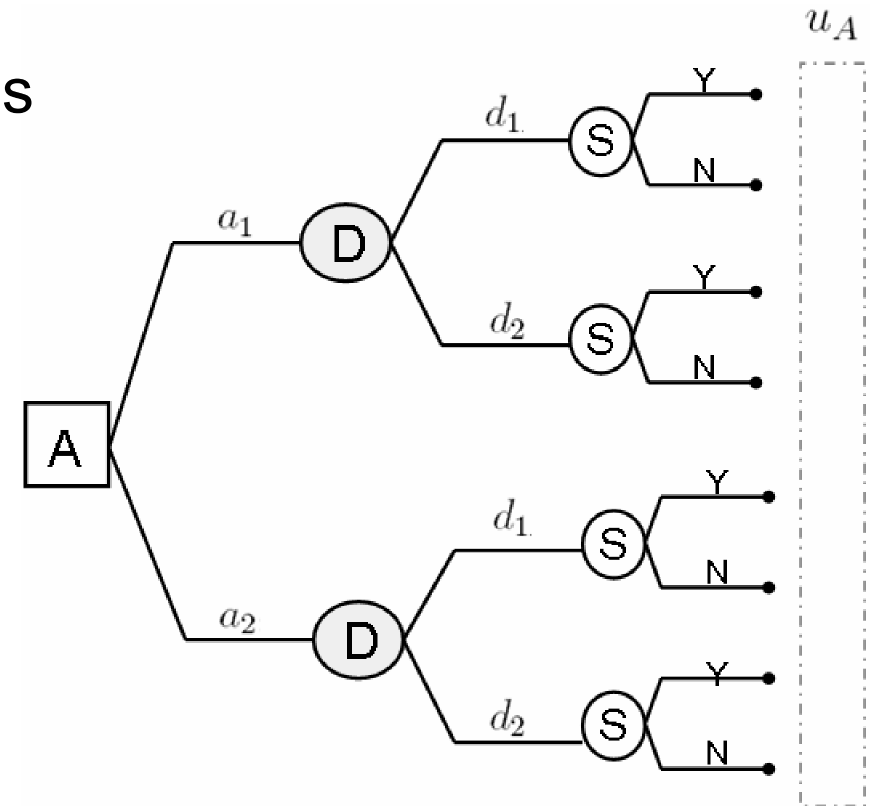
$$d^* = \operatorname{argmax}_{d \in \mathcal{D}} \sum_{a \in \mathcal{A}} \underbrace{\left[\sum_{s \in \{0,1\}} u_D(d,s)\, p_D(S = s \mid d, a) \right]}_{\psi_D(d,a)} \pi_D(A = a)$$

How to elicit it ??

Assessing: $\pi_D(A = a)$

- Attacker's decision analysis as seen by the Defender





$$a^* = \text{argmax}_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} \underbrace{\left[\sum_{s \in \{0,1\}} u_A(a, s)\; p_A(S = s \mid d, a) \right]}_{\psi_A(d,a)} \pi_A(D = d)$$

$$(u_A, p_A, \pi_A) \sim (U_A, P_A, \Pi_A)$$

Assessing $\pi_D(A = a)$

$$A \mid D \sim \operatorname{argmax}_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} \left[\underbrace{\sum_{s \in \{0,1\}} U_A(a, s)\ P_A(S = s \mid d, a)}_{\Psi_A(d, a)} \right] \Pi_A(D = d)$$

- $\Pi_A(D = d)$
  - Attacker's uncertainty about Defender's decision $\pi_A(D = d)$
  - Defender's uncertainty about the model used by the Attacker to predict what defense the Defender will choose $\pi_A \sim \Pi_A$

- The elicitation of $\Pi_A(D = d)$ may require further analysis
  Next level of recursive thinking

$$D \mid A^1 \sim \operatorname{argmax}_{d \in \mathcal{D}} \sum_{a \in \mathcal{A}} \left[\underbrace{\sum_{s \in \{0,1\}} U_D(d, s)\ P_D(S = s \mid d, a)}_{\Psi_D(d, a)} \right] \Pi_D(A^1 = a)$$

The assessment problem

- To predict Attacker's decision

  The Defender needs to solve Attacker's decision problem

  She needs to assess (u_A, p_A, π_A)

- Her beliefs about $(u_A, p_A, \pi_A) \sim (U_A, P_A, \Pi_A)$

- The assessment of $\Pi_A(D = d)$ requires further analysis
  - D's analysis of A's analysis of D's problem
    Thinking-about-what-the-other-is-thinking-about…

- It leads to a hierarchy of nested decision models

Hierarchy of nested decision models

Repeat

Find $\Pi_{D^{i-1}}(A^i)$ by solving

$$A^i \mid D^i \quad \sim \quad \text{argmax}_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} \left[\sum_{s \in \{0,1\}} U_A^i(a,s) \; P_A^i(S = s \mid d, a) \right] \Pi_{A^i}(D^i = d)$$

$$\text{where } (U_A^i, P_A^i) \sim F^i$$

Find $\Pi_{A^i}(D^i)$ by solving

$$D^i \mid A^{i+1} \quad \sim \quad \text{argmax}_{d \in \mathcal{D}} \sum_{a \in \mathcal{A}} \left[\sum_{s \in \{0,1\}} U_D^i(d,s) \; P_D^i(S = s \mid d, a) \right] \Pi_{D^i}(A^{i+1} = a)$$

$$\text{where } (U_D^i, P_D^i) \sim G^i$$

$i = i + 1$

Stop when the Defender has no more information about utilities and probabilities at some level of the recursive analysis

How to stop this infinite regress?

- Potentially infinite analysis of nested decision models

D → DA → DAD → DADA → DADAD → …

d* ← A ← D ← A^1 ← D^1 ← …

(u_D, p_D) (u_A, p_A)

(U_A, P_A) (U_D, P_D) (U_A^1, P_A^1) …

- Game Theory
  - Full and common knowledge assumption: $\begin{cases} d^* = \operatorname{argmax}_{d \in \mathcal{D}} \ \psi_D(d, a^*) \\ a^* = \operatorname{argmax}_{a \in \mathcal{A}} \ \psi_A(d^*, a) \end{cases}$

    (u_A, p_A, u_D, p_D)
  - Common prior assumption: $\begin{cases} A = A^1 = \ldots \\ D = D^1 = \ldots \end{cases}$

    (U_A, P_A, U_D, P_D)

- ARA: where to stop?
  - when no more info can be accommodated
  - Non-informative or reference model
  - Sensitivity analysis test

A numerical example

- Defender chooses d_1 or d_2
- Simultaneously Attacker must choose a_1 or a_2
- Defender assessments:

$u_D(d, s)$

	$s = 1$	$s = 0$
d_1	50	80
d_2	0	100

$p_D(S = 1 \mid d, a)$

	a_1	a_2
d_1	0.1	0
d_2	0.9	0

  – Two different types of Attacker
    - Type I prob 0.8
    - Type II prob 0.2

Skip example

$(U_{A_I}, P_{A_I}) \sim F_I$:

$U_{A_I}(a, s)$

	$s=1$	$s=0$
a_1	$Tri(20, 100, 100)$	$Tri(0, 20, 100)$
a_2	100	$Tri(0, 40, 100)$

$P_{A_I}(S=1 \mid d, a)$

	a_1	a_2
d_1	$\mathcal{U}[0, 1]$	0
d_2	$Tri(0.5, 1, 1)$	0

$(U_{A_{II}}, P_{A_{II}}) \sim F_{II}$:

$U_{A_{II}}(a, s)$

	$s=1$	$s=0$
a_1	$\mathcal{U}[0, 100]$	$Tri(0, 20, 100)$
a_2	100	$Tri(40, 80, 90)$

$P_{A_{II}}(S=1 \mid d, a)$

	a_1	a_2
d_1	$Tri(0, 0, 1)$	0
d_2	$Tri(0, 1, 1)$	0

- Defender thinks that a Type I Attacker is intelligent enough to analyze her problem
  - A Type I Attacker's beliefs about her utilities and probabilities are

$(U_{D_I}, P_{D_I}) \sim G_I$:

$U_{D_I}(d, s)$

	$s = 1$	$s = 0$
d_1	$Tri(0, 0, 40)$	$\mathcal{U}[50, 100]$
d_2	$Tri(0, 0, 40)$	$\mathcal{U}[50, 100]$

$P_{D_I}(S = 1 \mid d, a)$

	a_1	a_2
d_1	$Tri(0, 0, 0.5)$	0
d_2	$\mathcal{U}[0, 1]$	0

$$\Pi_{A_I}(D_I = d_1) \sim \mathcal{B}e(\alpha, 10 - \alpha), \text{ where } \alpha = \pi_{A_I}(D_I = d_1) \times 10$$

- However, the Defender does not know how a Type II Attacker would analyze her problem, but believes that

$$\Pi_{A_{II}}(D_{II} = d_1) \sim \mathcal{B}e(75, 25)$$

- Defender: what does Type I Attacker think to be her beliefs about what he will do?

$$\Pi_{D_I}(A_I^1 = a_1) \sim \mathcal{U}[0, 1]$$

- Solving Defender's decision problem
  - Computing her defense of max. expected utility
- She first needs to compute
  - Her predictive distribution about what an Attacker will do

$$\pi_D(A = a_1) = 0.8 \times \pi_D(A_I = a_1) + 0.2 \times \pi_D(A_{II} = a_1)$$

$\pi_D(A_I = a_1) \longrightarrow$

1. For $k = 1, \ldots, n$, repeat
   - Draw $\pi^k_{D_I} \sim \Pi_{D_I}$, that is $\pi^k_{D_I}(A^1_I = a_1) \sim \mathcal{U}[0,1]$.
   - Draw $(u^k_{D_I}, p^k_{D_I}) \sim (U_{D_I}, P_{D_I}) = G_I$
   - Compute

$$d^k_I = \text{argmax}_{d \in \mathcal{D}} \sum_{a \in \mathcal{A}} \left[\sum_{s \in \{0,1\}} u^k_{D_I}(d, s)\, p^k_{D_I}(S = s \mid d, a) \right] \pi^k_{D_I}(A^1_I = a)$$

2. Approximate $\pi_{A_I}(D_I = d_1)$ through $\hat{\pi}_{A_I}(D_I = d_1) = \#\{d^k_I = d_1\}/n$.

   Set $\hat{\Pi}_{A_I}(D_I = d_1) \sim \mathcal{B}e(\alpha, 10 - \alpha)$, with $\alpha = \hat{\pi}_{A_I}(D_I = d_1) \times 10$.

3. For $k = 1, \ldots, n$, repeat
   - Draw $\hat{\pi}^k_{A_I} \sim \hat{\Pi}_{A_I}$, that is $\hat{\pi}^k_{A_I}(D_I = d_1) \sim \hat{\Pi}_{A_I}(D_I = d_1)$
   - Draw $(u^k_{A_I}, p^k_{A_I}) \sim (U_{A_I}, P_{A_I}) = F_I$
   - Compute

$$a^k_I = \text{argmax}_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} \left[\sum_{s \in \{0,1\}} u^k_{A_I}(a, s)\, p^k_{A_I}(S = s \mid d, a) \right] \hat{\pi}^k_{A_I}(D_I = d)$$

4. Approximate $\pi_D(A_I = a_1)$ through $\hat{\pi}_D(A_I = a_1) = \#\{a^k_I = a_1\}/n$.

$\pi_D(A_{II} = a_1) \longrightarrow$

1. For $k = 1, \ldots, n$, repeat
   - Draw $\pi^k_{A_{II}} \sim \Pi_{A_{II}}$, that is $\pi^k_{A_{II}}(D_{II} = d_1) \sim \mathcal{B}e(75, 25)$.
   - Draw $(u^k_{A_{II}}, p^k_{A_{II}}) \sim (U_{A_{II}}, P_{A_{II}}) = F_{II}$
   - Compute

$$a^k_{II} = \text{argmax}_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} \left[\sum_{s \in \{0,1\}} u^k_{A_{II}}(a, s)\, p^k_{A_{II}}(S = s \mid d, a) \right] \pi^k_{A_{II}}(D_{II} = d)$$

2. Approximate $\pi_D(A_{II} = a_1)$ through $\hat{\pi}_D(A_{II} = a_1) = \#\{a^k_{II} = a_1\}/n$.

– In a run with n=1000, we got

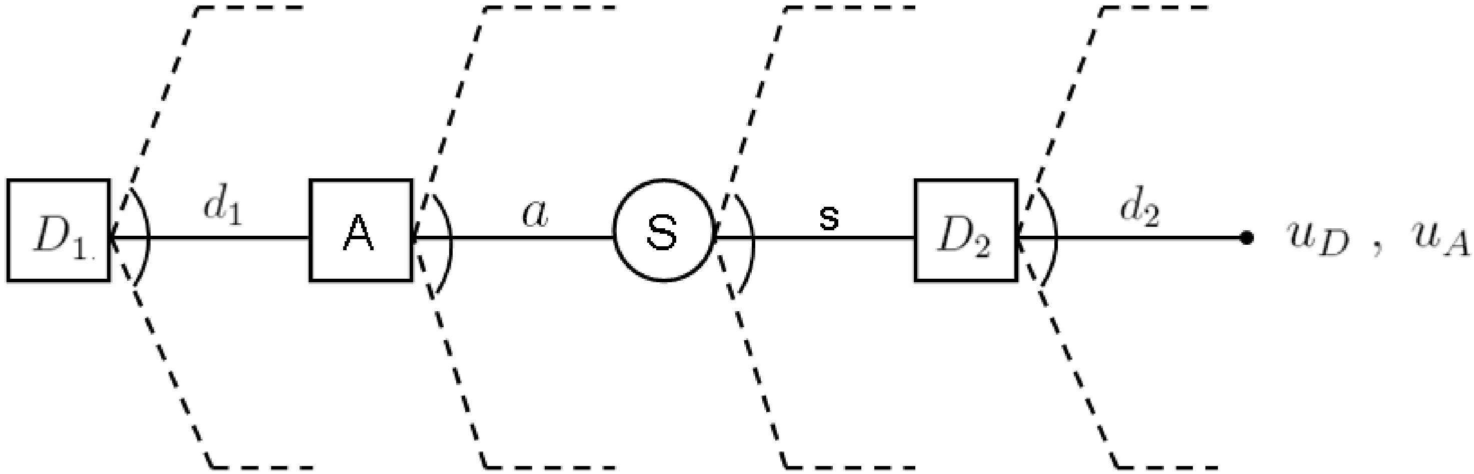
$$\hat{\pi}_D(A_I = a_1) = 0.97 \quad \times \quad 0.8$$
$$\hat{\pi}_D(A_{II} = a_1) = 0.82 \quad \times \quad 0.2$$
$$\hat{\pi}_D(A = a_1) = 0.94$$

- And, now the Defender can solve her problem

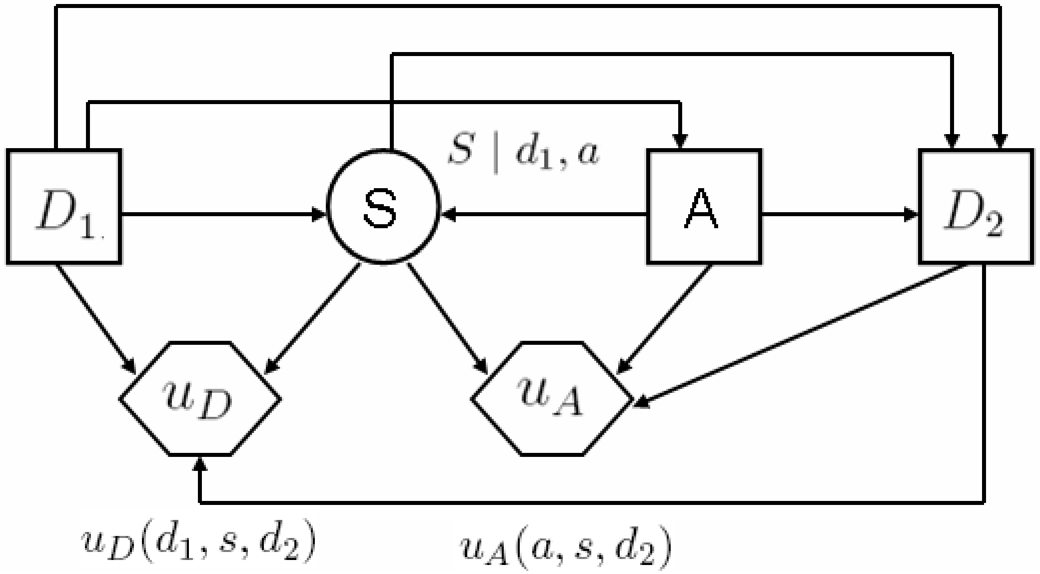
$$d^* = \text{argmax}_{d \in \mathcal{D}} \sum_{a \in \mathcal{A}} \left[\sum_{s \in \{0,1\}} u_D(d, s)\, p_D(S = s \mid d, a) \right] \pi_D(A = a)$$

$d^* = d_1$ with (MC estimated) expected utility 77, against d_2 with 15

Defend–Attack–Defend model



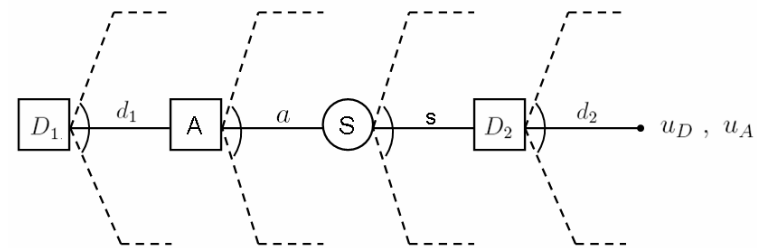
skip



Standard Game Theory Analysis

- Under common knowledge of utilities and probs
- At node D_2

$$d_2^*(d_1, s) = \text{argmax}_{d_2 \in \mathcal{D}_2} u_D(d_1, s, d_2)$$



- Expected utilities at node S

$$\psi_D(d_1, a) = \int u_D(d_1, s, d_2^*(d_1, s))\, p_D(s \mid d_1, a)\, \mathrm{d}s$$

$$\psi_A(d_1, a) = \int u_A(a, s, d_2^*(d_1, s))\, p_A(s \mid d_1, a)\, \mathrm{d}s$$

- Best Attacker's decision at node A

$$a^*(d_1) = \text{argmax}_{a \in \mathcal{A}} \psi_A(d_1, a)$$

- Best Defender's decision at node D_1

$$d_1^* = \text{argmax}_{d_1 \in \mathcal{D}_1} \psi_D(d_1, a^*(d_1))$$

- Nash Solution: $d_1^* \in \mathcal{D}_1 \quad a^*(d_1^*) \in \mathcal{A} \quad d_2^*(d_1^*, s) \in \mathcal{D}_2$

ARA: Supporting the Defender

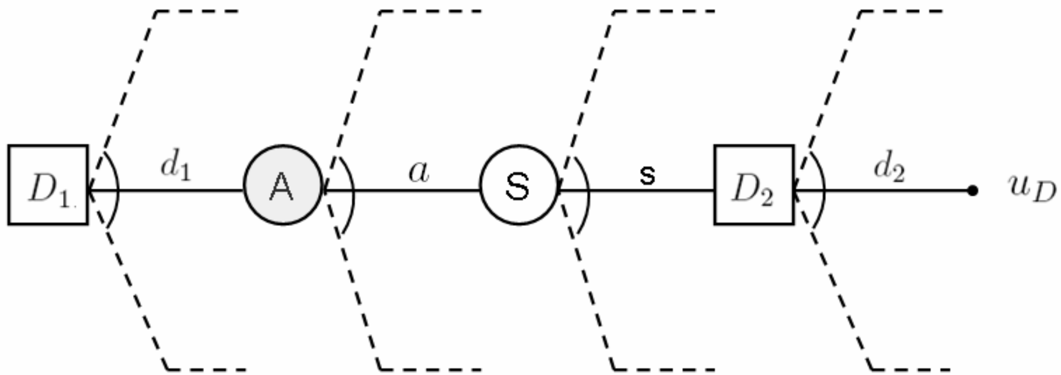
- At node A

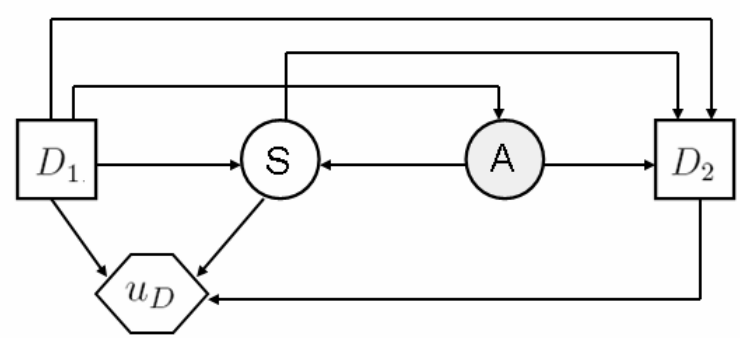
$$\psi_D(d_1) = \int \psi_A(d_1, a)\, p_D(a \mid d_1)\, \mathrm{d}a$$

- At node D_1

$$d_1^* = \mathrm{argmax}_{d_1 \in \mathcal{D}_1} \psi_D(d_1)$$

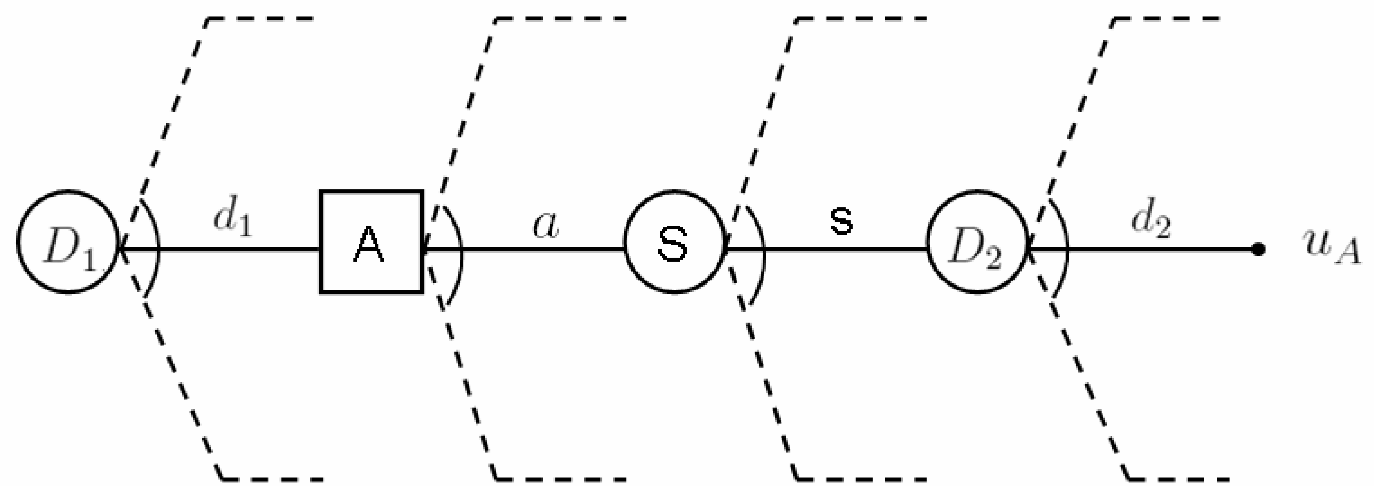
- $p_D(A \mid d_1)$??

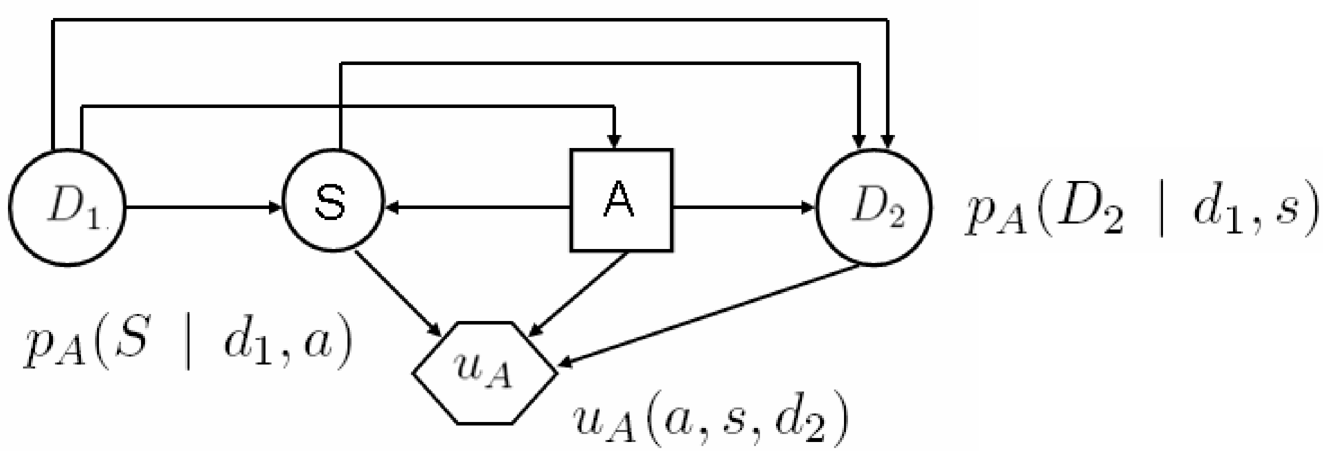




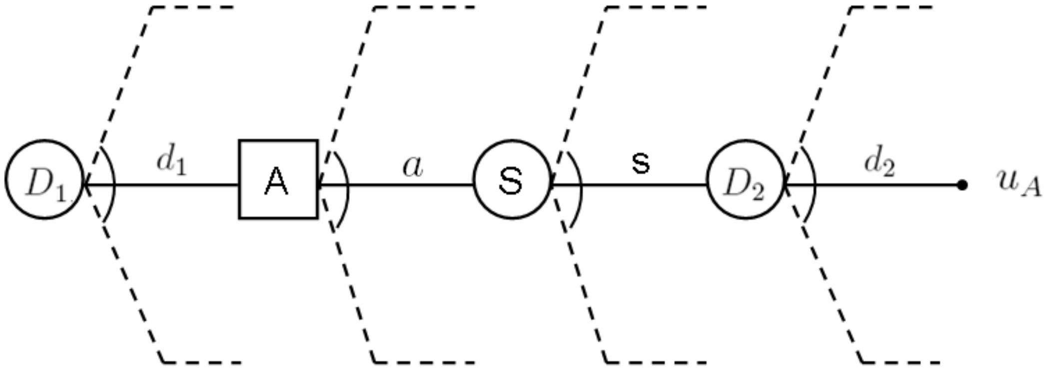
Assessing $p_D(A \mid d_1)$

- Attacker's problem as seen by the Defender





Assessing $p_D(A \mid d_1)$



- At chance node D_2, compute

$$(d_1, a, s) \to \Psi_A(d_1, a, s) = \int U_A(a, s, d_2)\ P_A(D_2 = d_2 \mid d_1, s)\ \mathrm{d}d_2$$

- At chance node S

$$(d_1, a) \to \Psi_A(d_1, a) = \int \Psi_A(d_1, a, s)\ P_A(S = s \mid d_1, a)\ \mathrm{d}s$$

- At decision node A

$$d_1 \to A^*(d_1) = \mathrm{argmax}_{a \in \mathcal{A}} \Psi_A(d_1, a)$$

- $p_D(A = a \mid d_1) = \Pr(A^*(d_1) = a)$

Monte-Carlo approximation of $p_D(A \mid d_1)$

- Drawn $\{(u_A^i(a,s,d_2), p_A^i(S \mid d_1, a), p_A^i(D_2 \mid d_1, s))\}_{i=1}^n \sim F$

- Generate $\{a_i^*(d_1)\}_{i=1}^n$ by

  - At chance node D_2

    $$(d_1, a, s) \to \psi_A^i(d_1, a, s) = \int u_A^i(a, s, d_2)\; p_A^i(D_2 = d_2 \mid d_1, s)\; \mathrm{d}d_2$$

  - At chance node S

    $$(d_1, a) \to \psi_A^i(d_1, a) = \int \psi_A^i(d_1, a, s)\; p_A^i(S = s \mid d_1, a)\; \mathrm{d}s$$

  - At decision node A

    $$d_1 \to a_i^*(d_1) = \mathrm{argmax}_{a \in \mathcal{A}} \psi_A^i(d_1, a)$$

- Approximate

$$p_D(A = a \mid d_1) \approx \#\{a_i^*(d) = a\}/n$$

The assessment of $p_A(D_2 \mid d_1, s)$

- The Defender may want to exploit information about how the Attacker analyzes her problem

- Hierarchy of recursive analysis

Discussion

- DA vs GT
  - A Bayesian prescriptive approach to support a Defender against an Attacker
    - Computation of her defense of maximum expected utility
  - Weaken common (prior) knowledge assumption
  - Analysis and assessment of Attacker' thinking to anticipate his actions
    - The assessment problem under infinite regress
- We have assumed that the Attacker is a expected utility maximizer
  - Other *descriptive* models of rationality (non expected utility models)
- Several simple but illustrative models
  - What if
    - more complex dynamic interactions?
    - against more than one Attacker or an uncertain number of them?
- More than one agent at each side
  - Two or more countries coordinate resources to counter two or more terrorist groups
  - External model on the intelligent adversaries' behaviour
- Implementation issues
  - Elicitation of a valuable judgmental input from Defender
  - Computational issues
- Real problems

Some references

- Banks, D. and S. Anderson (2006) Game theory and risk analysis in the context of the smallpox threat, in A. Wilson, G. Wilson and D. Olwell (ed) *Statistical Methods in Counterterrorism*, 9-22.
- Kadane, J.B. and P.D. Larkey (1982) Subjective probability and the theory of games, *Management Science*, 28, 113-120.
- Parnell, G. (2007) Multi-objective Decision Analysis, in Voeller (ed) *Handbook of Science and Technology for Homeland Security*, Wiley.
- Parnell, G., Banks, D., Borio, L., Brown, G., Cox, L. A., Gannon, J., Harvill, E., Kunreuther, H., Morse, S., Pappaioanou, M., Pollack, S., Singpurwalla, N., and Wilson, A. (2008). Report on Methodological Improvements to the Department of Homeland Security's Biological Agent Risk Analysis, National Academies Press.
- Pate-Cornell, E. and S. Guikema (2002) Probabilistic modeling or terrorist threats: a systematic analysis approach to setting priorities among countermeasures, *Military Operations Research*, 7, 5-23.
- Raiffa, H. (2002) *Negotiation Analysis*, Harvard University Press.
- Rios Insua, D. J. Rios, and D. Banks (2009) Adversarial risk analysis, Journal of the American Statistical Association, 104, 841-854.
- von Winterfeldt, D. and T.M. O'Sullivan (2006) Should we protect commercial airplanes against surface-to-air missile attacks by terrorists? *Decision Analysis*, 3, 63-75.