CISAC
Stanford

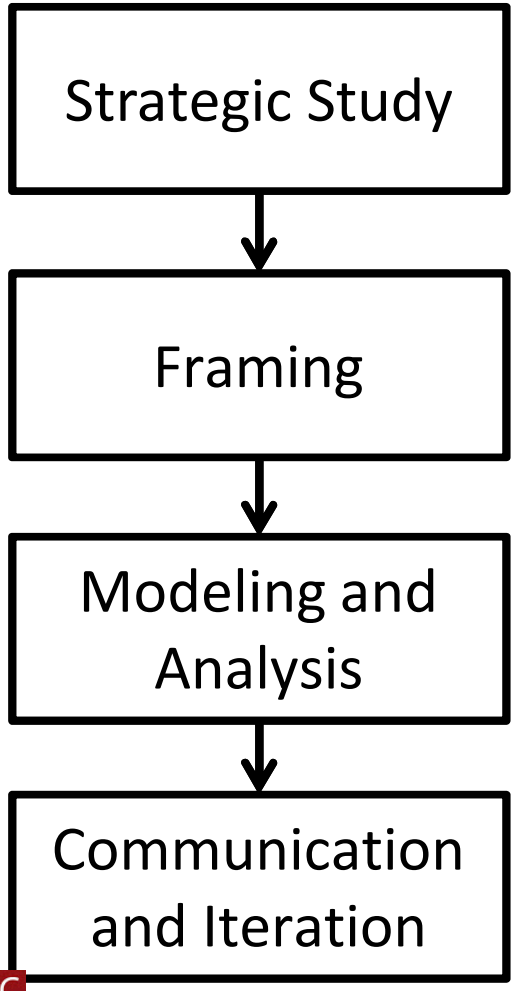
SAND2016-8134PE

Modeling Adversaries: A Survey of Methods

Jason Reinhardt
reinhardt@stanford.edu

Stanford Center at Peking University (SCPKU)
June 6-7, 2016

Phases of Systems Analysis

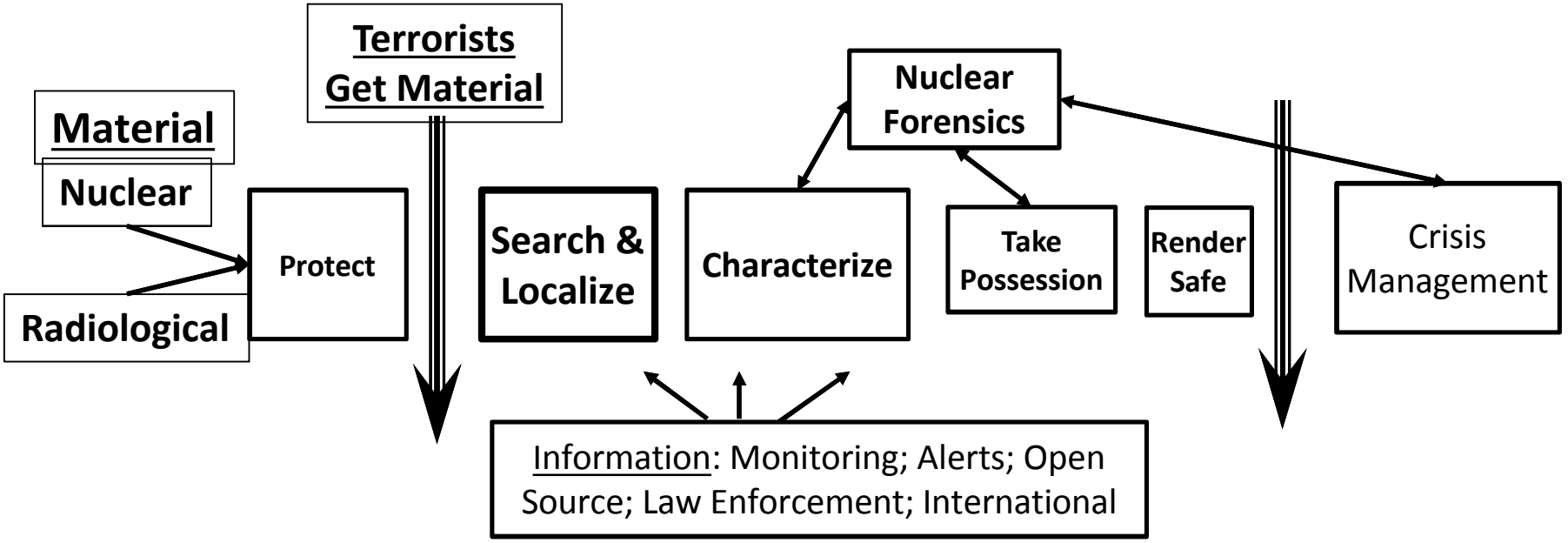


	Questions Addressed	**Examples**
Strategic Study	What is important? What is the problem?	Forensics (Brandt)
Framing	What is the scope? What are the alternatives?	Transit Ports (Zhang)
Modeling and Analysis	How do we systematically assess the alternatives?	Wide Area Search (Reinhardt) Adversary Models (Reinhardt)
Communication and Iteration	How do we communicate results? What is the impact?	Radiological (Connell)

Outline

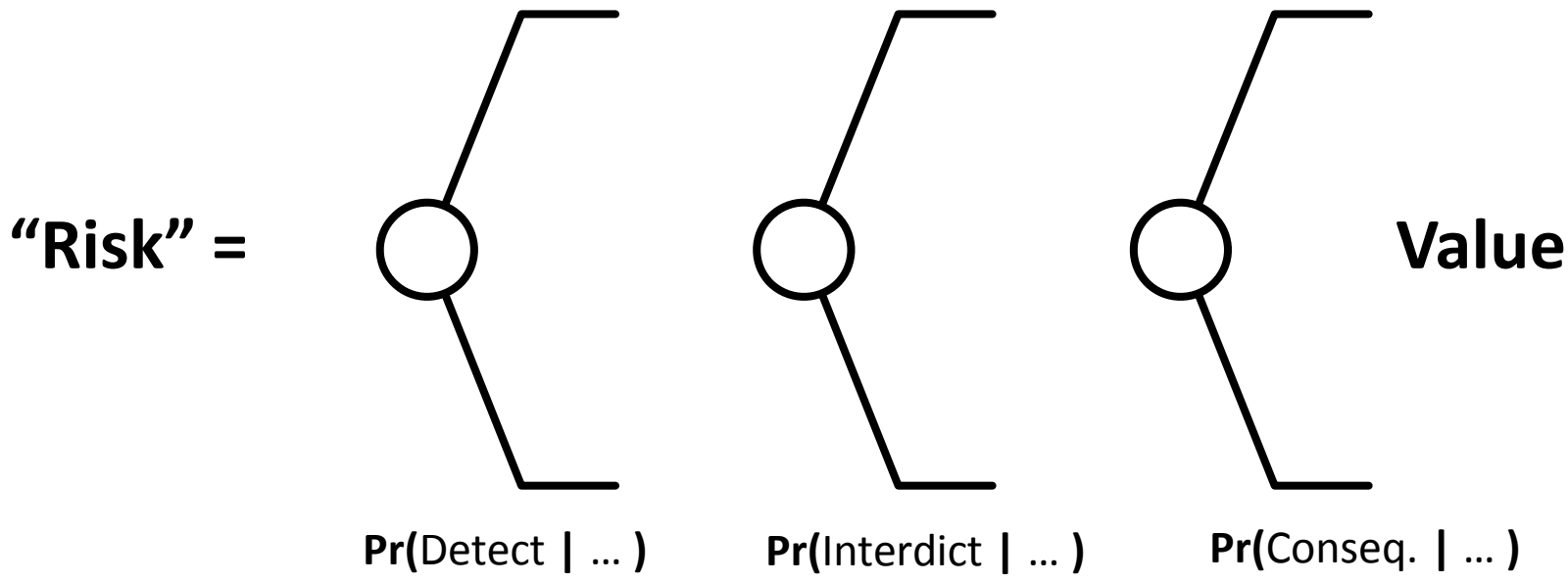
- Motivation
- Possible Approaches
  - Non-probability Based
  - Probability or Game Based
- Conclusions

End-to-end performance metrics are important in security efforts.



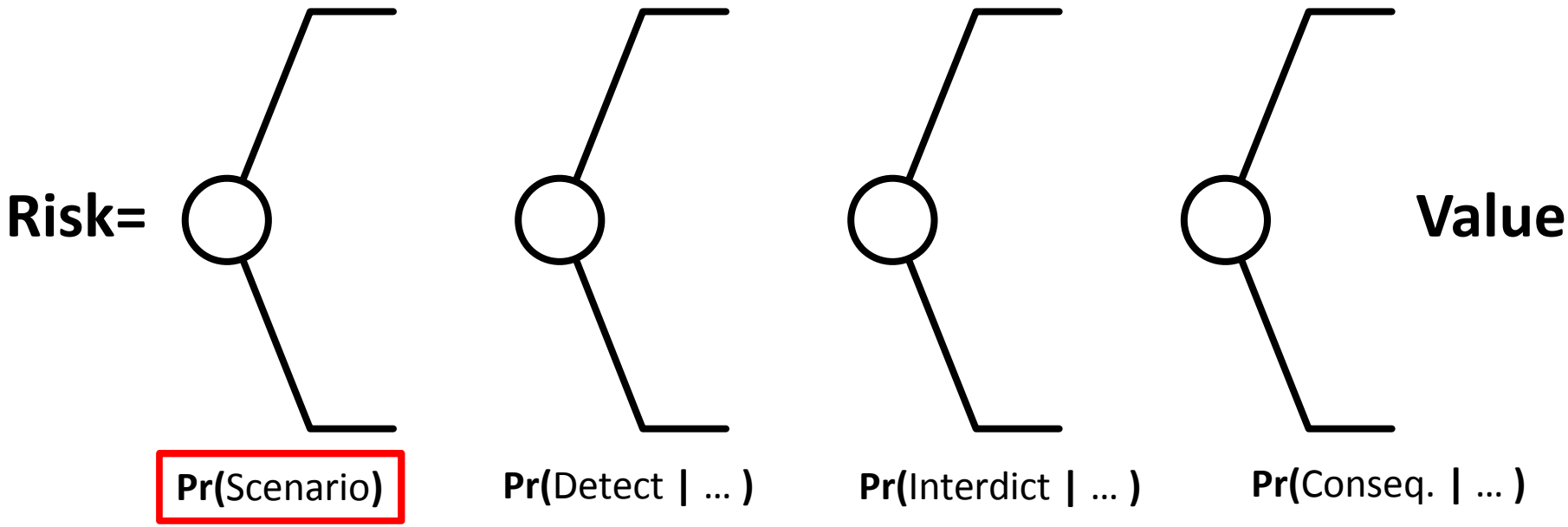
- Sets priorities for effort and investment
- Informs design of system and components
- Provides a basis for evaluating effectiveness

The Most Common "Risk" Approach



- Can be done qualitatively or quantitatively
- Only provides an estimate of risk conditioned on the assumption that the adversary WILL attack
- Does not allow for deterrence, threat shifting, or other effects of strategic interaction

A full risk analytic approach focuses the adversary modeling problem.



- The adversary chooses the scenario:
  - Whether to attack, When to attack, How to attack
- Adversary choices influence the whole model structure
- Requires to estimate some form of probability distribution over the scenario space

CISAC Stanford

The Common Methods

- **Non-Probability Based**
  - **Red Team**
  - ***Adversary Capability Level (ACL)***
  - **Systematic Difficulty**

- Probability Based
  - Direct Elicitation
  - Empirical Assessment
  - Game Theoretic

The Red Team (or Blue Team?)

- Assumptions
  - ‘Red Team’ thinks like adversary
  - Scenario examined is likely
- Pros
  - Rich adversary model, creative, adaptive, experienced
- Cons
  - Expensive to run (little to no sensitivity analysis)
  - Difficult to replicate results
  - Difficult to be systematic

Can be useful for exposing issues and vulnerabilities, but conclusions regarding system performance are generally suggestive at best.

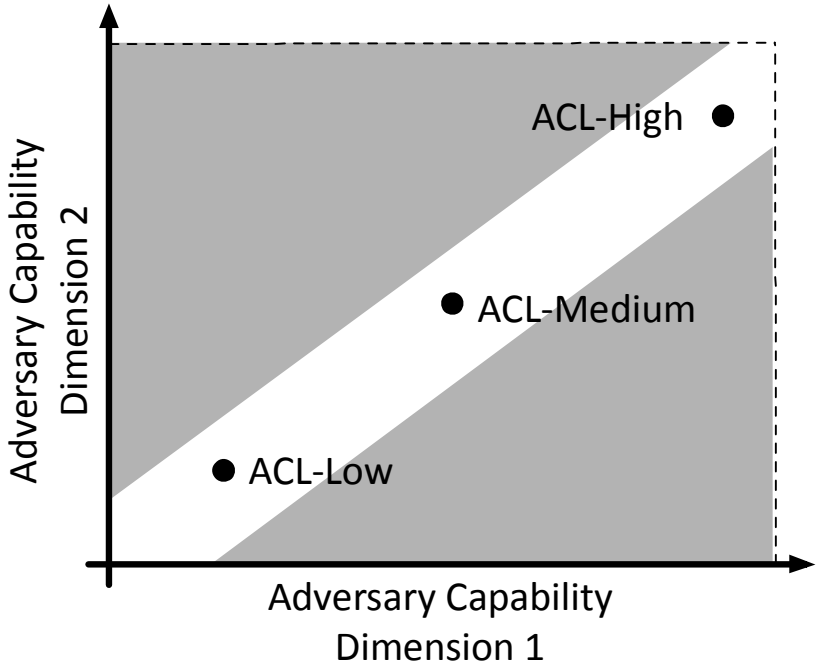
Adversary Capability Level (ACL)

Attributes	ACL-Low	ACL-Medium	ACL-High
General Description	Domestic Terrorist Group	Regional Terrorist Group	International Terrorist Group
Team Size	L	M	H
Funding Level	L	M	H
Critical Skill Sets	L	M	H
Critical Knowledge Sets	L	M	H

- Create adversary “types” that span a set of possible adversaries by focusing on varying levels of capability
- Can vary multiple kinds of capability
- Use those types to create scenarios for analysis
- Use resulting scenarios to assess alternate detection architecture and operations strategies
- **Method of choice for this workshop**

Adversary Capability Level (ACL)

- Assumptions
  - ACL classes sufficiently span possible set of adversaries
  - ACL analysts think like adversaries
- Pros
  - Examines multiple adversary "types"
  - Provides a consistent scenario basis set
  - Can indicate areas for deeper analysis and follow-on assessment



- Cons
  - May not represent dependencies between scenarios and mitigations accurately
  - Cannot easily correct for biases and bad intuition
  - Generally less detailed than red teaming

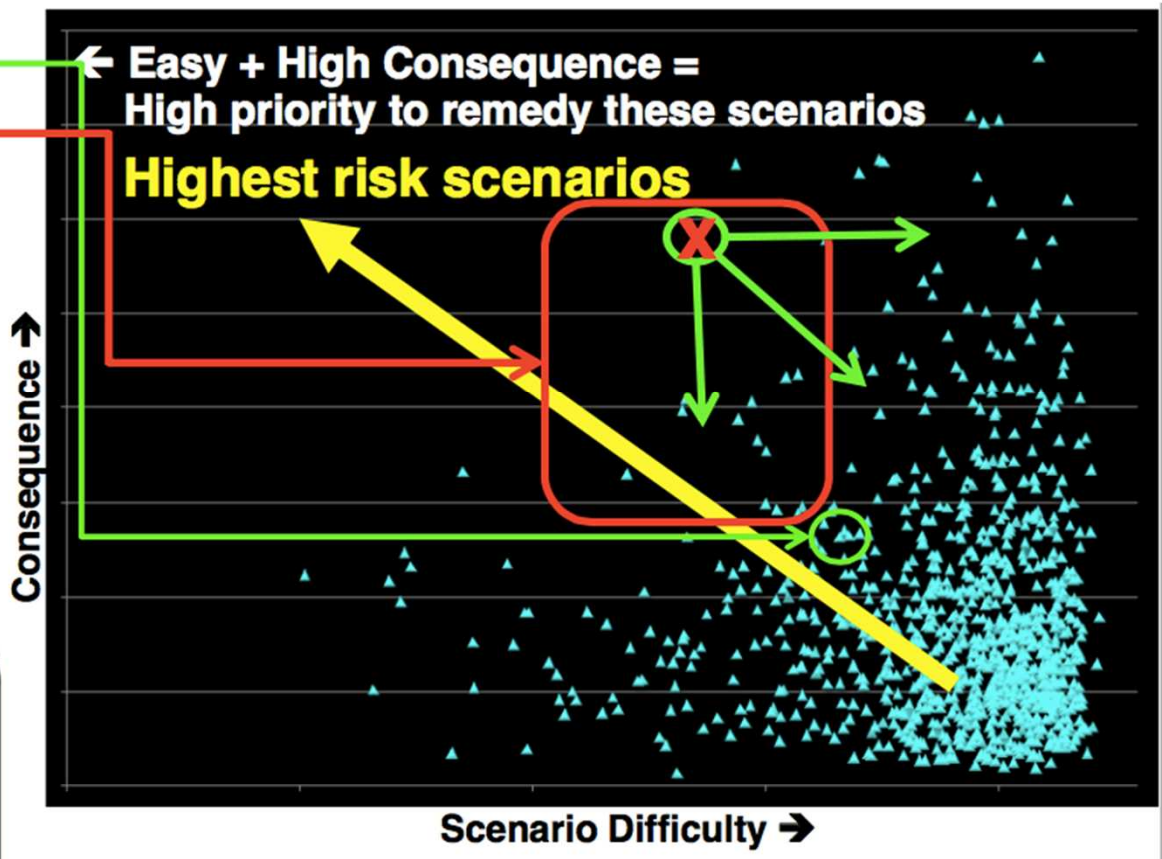
Systematic Difficulty: Approach

If we fix this…

Without fixing this…

We may not have improved security. *Because…*

Many scenarios still exist that are both easier to achieve AND provide higher consequences!



Why use scenario difficulty in security risk management?

- Difficulty better reflects the adversary planning process
- Difficulty changes more slowly and predictably than likelihood
- We have developed a qualitative (semi-quantitative) method to rank attack scenario difficulty

To "fix" a scenario we must

- Eliminate it (make it impossible to achieve)
- Reduce the consequences if it is completed
- Make it harder to accomplish successfully

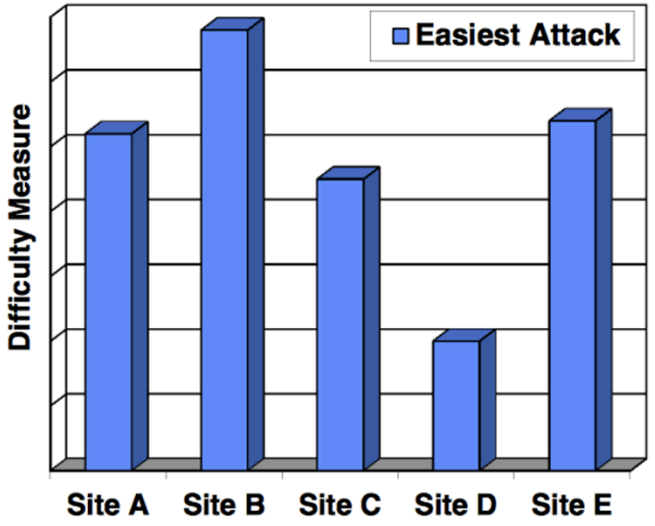
… or any combination of these

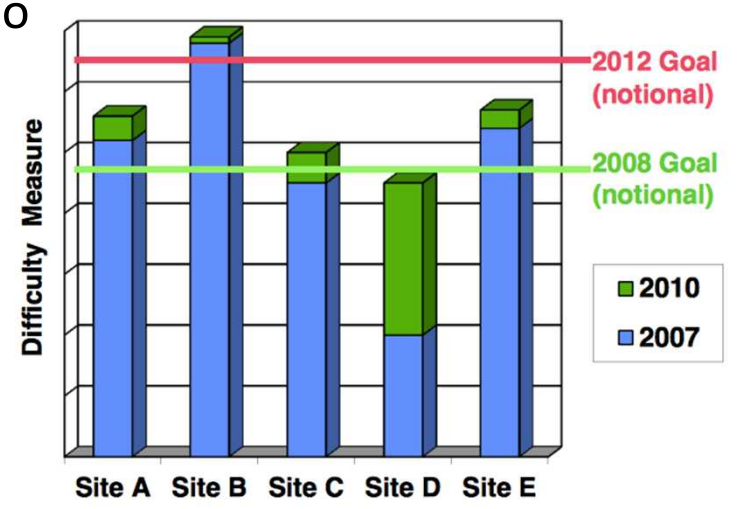
CISAC Stanford

Slide taken from: Wyss, "Risk-Informed Management of Enterprise Security…"

Systematic Difficulty: Assessment

- Assumptions
  - Adversaries are **less likely** to choose attack scenarios that are **more difficult** to accomplish
  - Defenders will prioritize attack scenarios that have more damaging potential consequences
- Pros
  - Focuses on qualitatively prioritizing risk management options, rather than assessing probabilistic modeling
  - Formalizes expert knowledge and opinion into a systematic framework
- Cons
  - May not represent dependencies between scenarios and mitigations accurately
  - Cannot easily correct for biases and bad intuition
  - May require input from many experts before results can be useful





CISAC Stanford

Images taken from: Wyss, "Risk-Informed Management of Enterprise Security…"

Systematic Difficulty: References

1. Gregory Wyss, “Risk-Informed Management of Enterprise Security: Method and Example Applications,” Sandia National Laboratories, Presentation to Institute for Nuclear Materials Management (INMM), http://www.inmm.org/AM/Template.cfm?Section=Risk_Informed_Security_Workshop1&Template=/CM/ContentDisplay.cfm&ContentID=4619

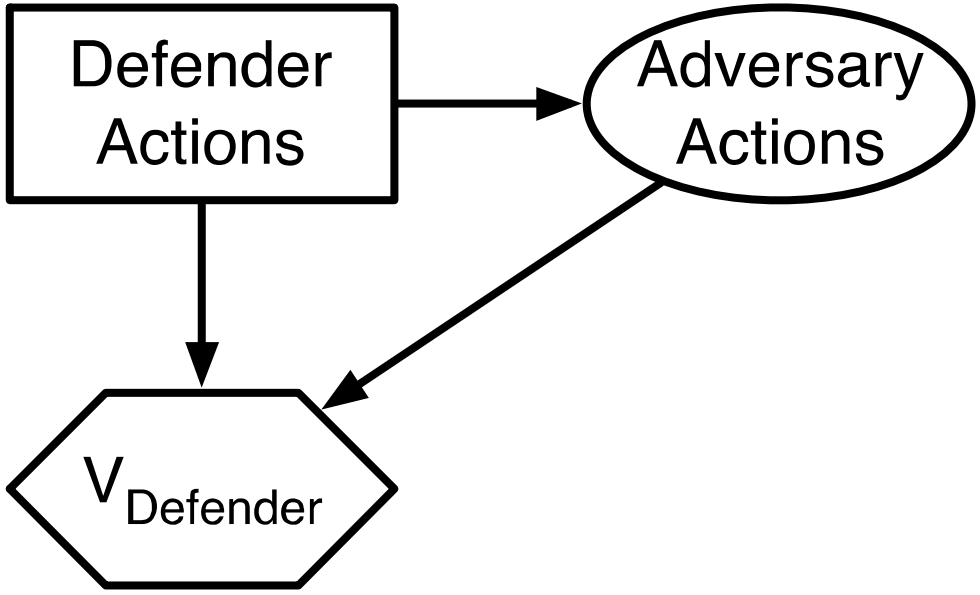
2. Gregory Wyss, “Risk-Based Cost-Benefit Analysis: Method and Example Applications,” Sandia National Laboratories, Presented at the INCOSE Enchantment Chapter Member Meeting, November 2011, http://www.incose.org/docs/default-source/enchantment/111109_gregorywyss_cost-benefit-slides.pdf?sfvrsn=2

The Common Methods

- Non-Probability Based
  - Red Team
  - *Adversary Capability Levels (ACL)*
  - Systematic Difficulty

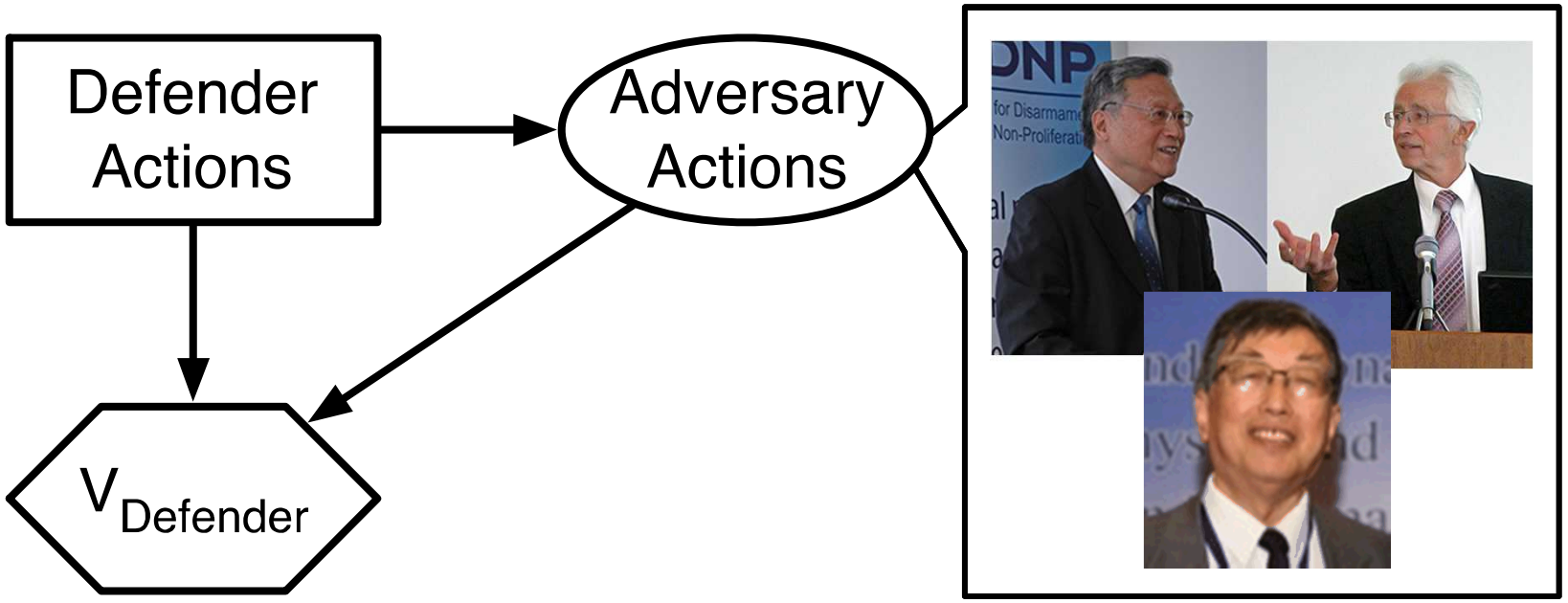
- **Probability or Game Based**
  - **Direct Elicitation**
  - **Empirical Analysis**
  - ***Game Theoretic***

Direct Elicitation: Ask an Expert!



- Relies on expert inputs on adversary characteristics, reactions, and decisions
- Requires careful elicitation techniques to account for biases
- Formalizes expert inputs in the framework of probability
- Utilizes risk and decision analysis approaches to inform architecture decisions

Direct Elicitation: Ask an Expert!

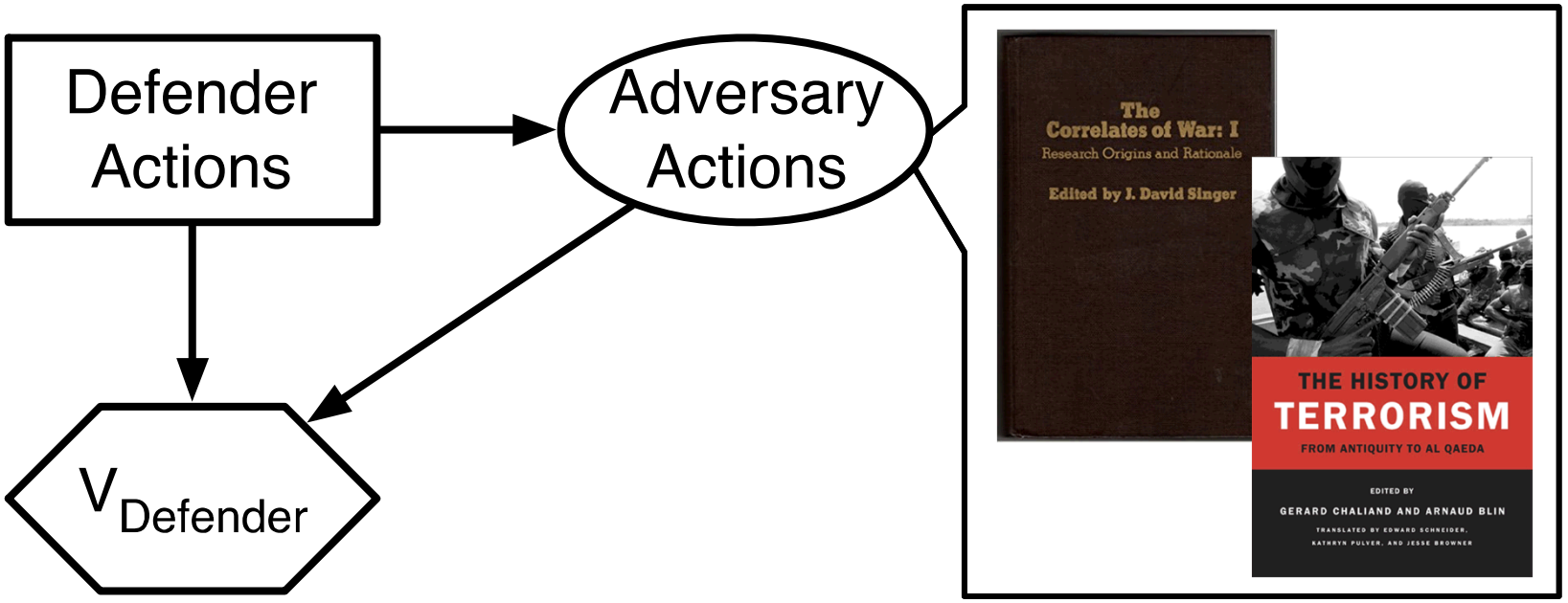


- Relies on expert inputs on adversary characteristics, reactions, and decisions
- Requires careful elicitation techniques to account for biases
- Formalizes expert inputs in the framework of probability
- Utilizes risk and decision analysis approaches to inform architecture decisions

Direct Elicitation: Assessment

- Assumptions
  - Experts understand adversaries well
  - Adversaries do not adapt, or adapt in simple ways
- Pros
  - Allows for the formulation of quantitative model
  - Enables early phase sensitivity and trade-off analysis
  - Most useful when combined with physics models
- Cons
  - Highly dependent on expert opinions, and can include their biases
  - Repeated elicitations may be required as analysis proceeds
  - Costly to perform assessments over large sets of possible scenarios

Empirical Analysis



- Collect data from historical case studies
- Assess probabilities for adversary actions based on statistical analysis of data
- Use those probabilities in analysis of optimum defender actions

Empirical Analysis: Assessment

- Assumptions
  - Past predicts future
  - Adversaries do not adapt
- Pros
  - Can sometimes provide a reasonable “prior”
  - Case studies can be illustrative
  - Enables early phase sensitivity and trade-off analysis
  - Most useful when combined with physics models
- Cons
  - Data often doesn’t exist, must use proxies
  - Costly to perform historical analysis over large sets of possible scenarios

It is important to remember the difference between two things:

Behavior of Groups

- Aggregate behaviors
- Markets, crime, politics
- "What is the probability that SOME person will…?"

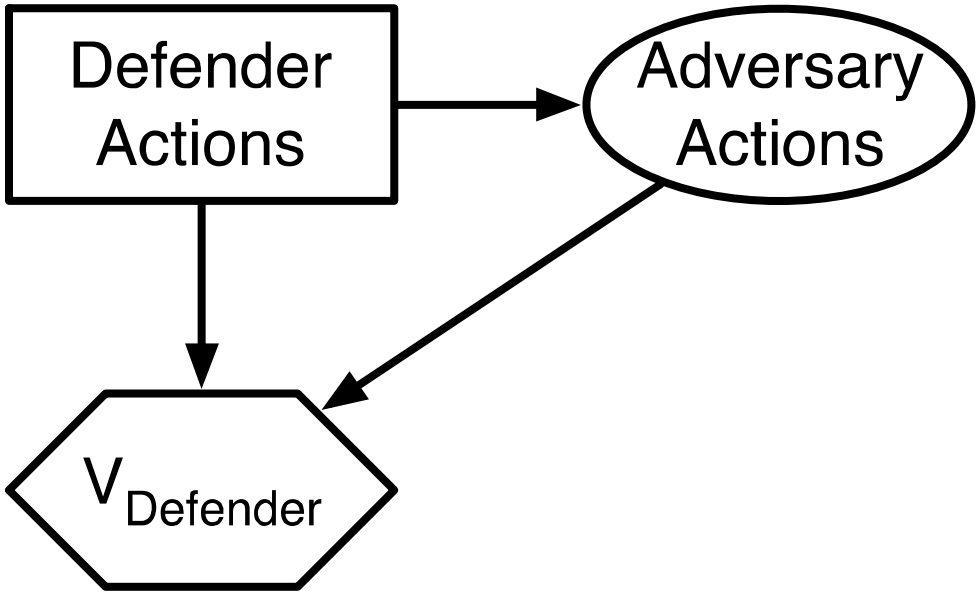
Behavior of Individuals

- Choices in strategic interaction and conflict
- "What is the probability that THIS person will…?"

CISAC
Stanford

Building an Illustrative Example

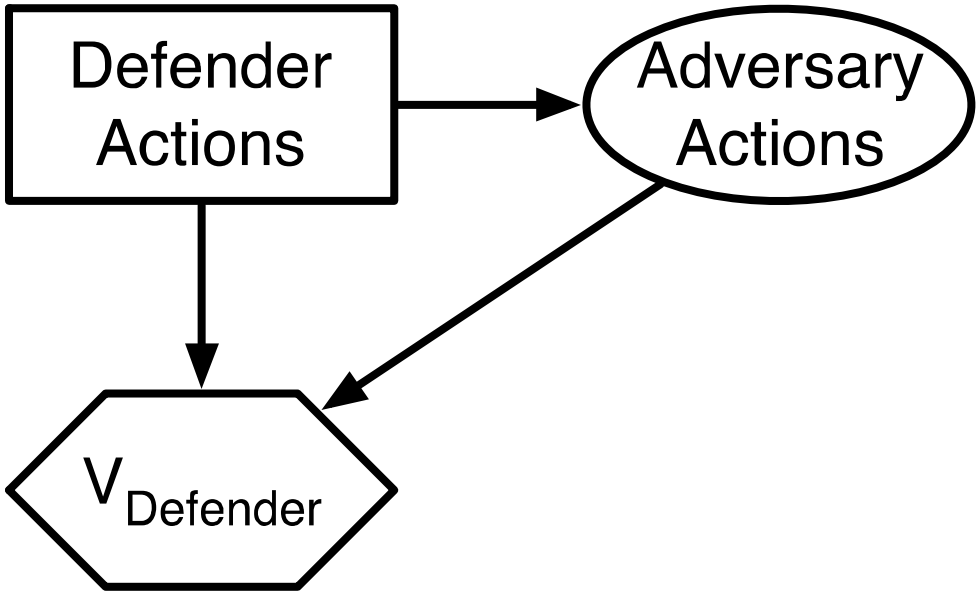
Defender chooses between:
Architecture A (e.g. Patrols)
Architecture B (e.g. Detectors)



Building an Illustrative Example

Defender chooses between:
Architecture A (e.g. Patrols)
Architecture B (e.g. Detectors)

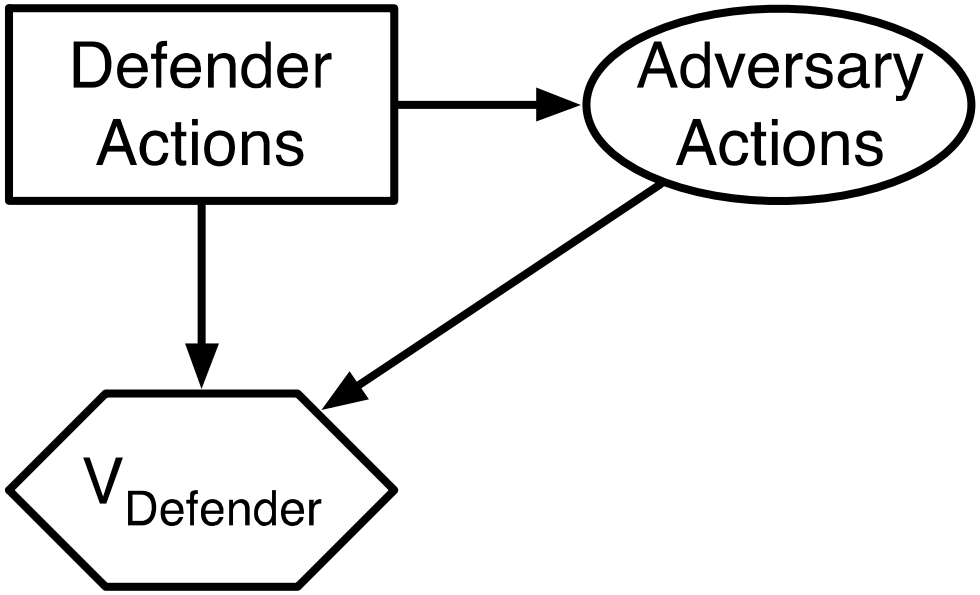
Adversary observes Defenders Actions, and then picks between:
Attack 1 (e.g. Maritime)
Attack 2 (e.g. Land)



Building an Illustrative Example

Defender chooses between:
Architecture A (e.g. Patrols)
Architecture B (e.g. Detectors)

Adversary observes Defenders Actions, and then picks between:
Attack 1 (e.g. Maritime)
Attack 2 (e.g. Land)

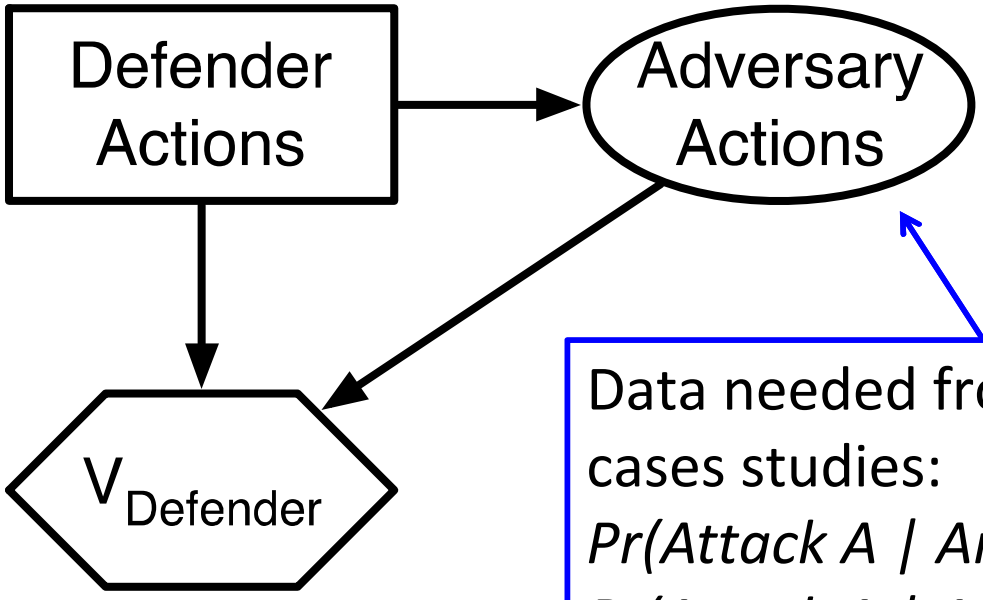


Value for Defender is assessed
as expected losses in each combination
of architecture and adversary actions.

Building an Illustrative Example

Defender chooses between:
Architecture A (e.g. Patrols)
Architecture B (e.g. Detectors)

Adversary observes Defenders Actions, and then picks between:
Attack 1 (e.g. Maritime)
Attack 2 (e.g. Land)

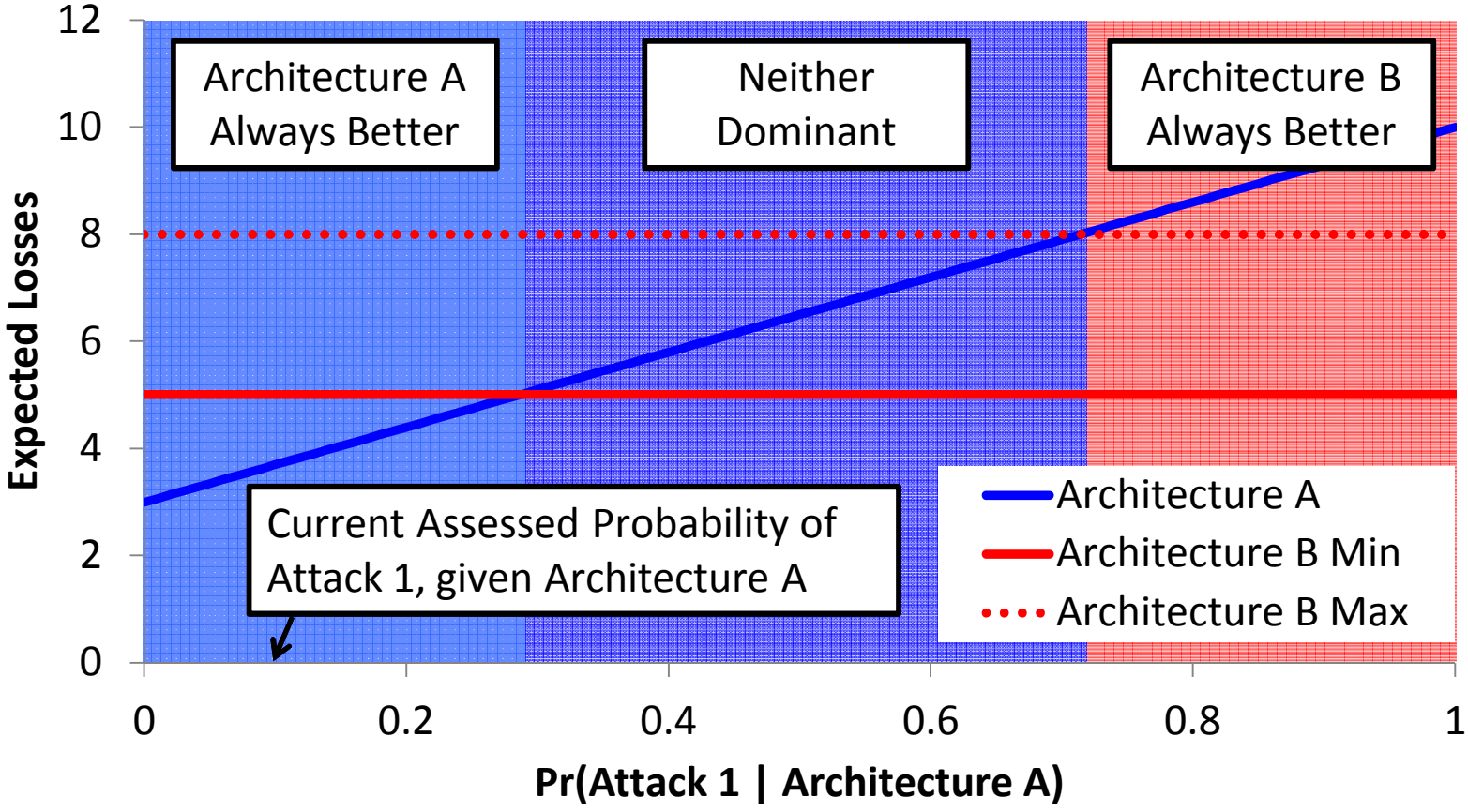


Data needed from experts or cases studies:
Pr(Attack A | Architecture A)
Pr(Attack A | Architecture B)

Value for Defender is assessed as expected losses in each combination of architecture and adversary actions.

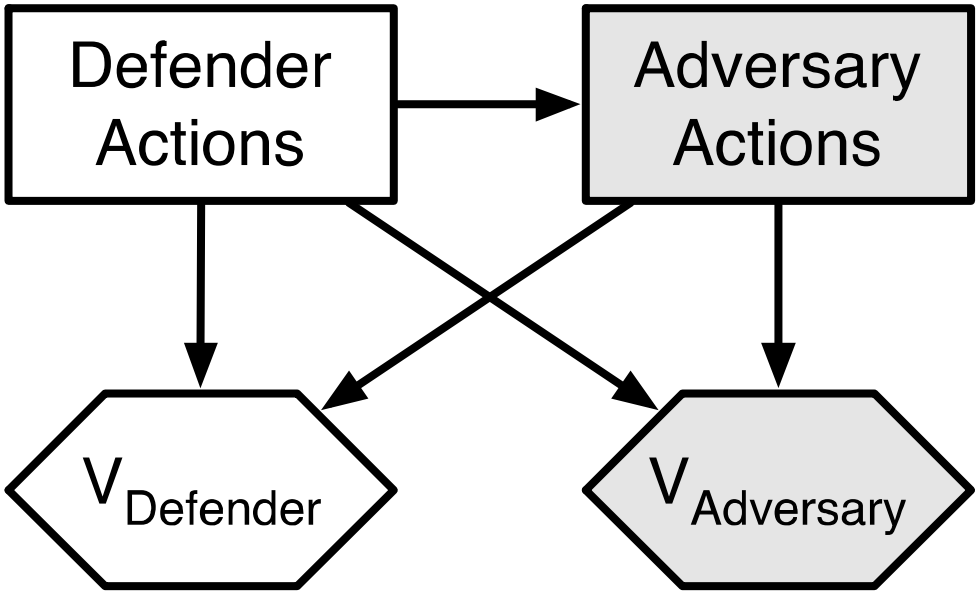
Using probabilities for an adversary model can reveal some insights.

Even if the expert doubles or triples their probability that the adversary will choose Attack 1 given Architecture A, the best option does not change. It is a stable choice, all else being equal.



Game Theoretic Methods

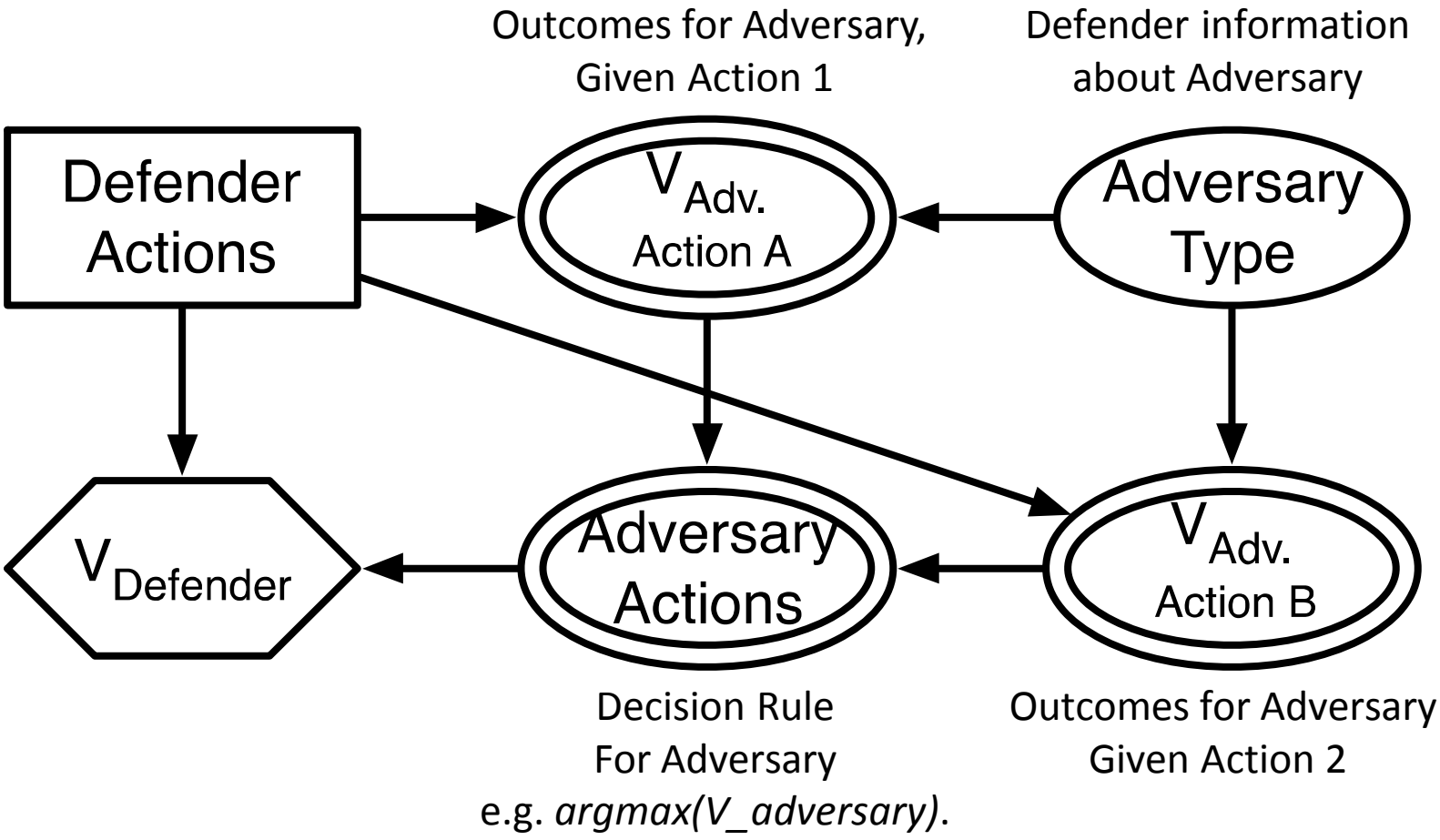
Model the decisions of the defender AND the decisions of the adversary.



But, how do we model the adversary's decisions?

Game Theoretic Method Models

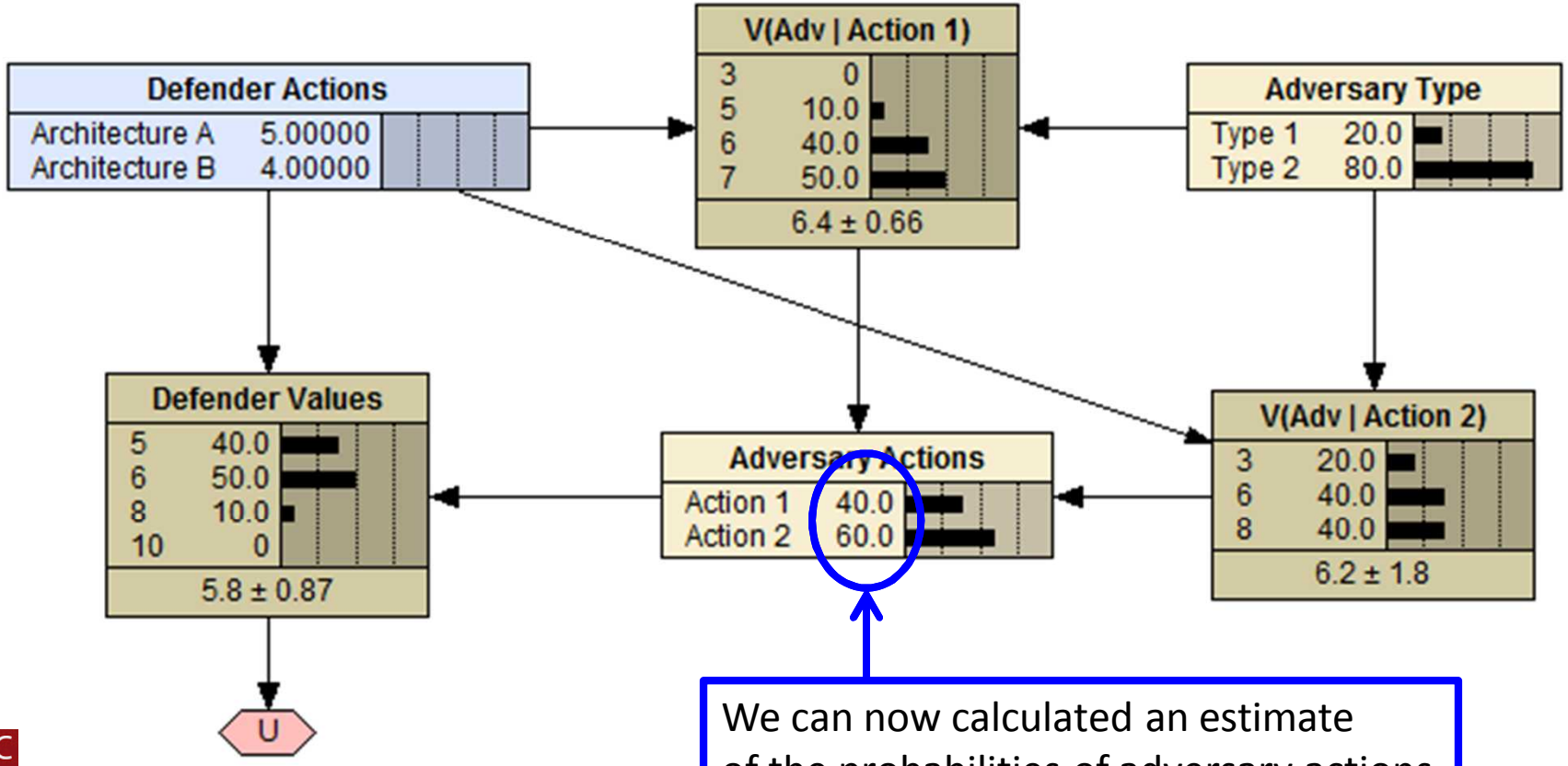
The "Adversarial Risk Analysis" technique systematically constructs an adversary model.



Game Theoretic Method Models

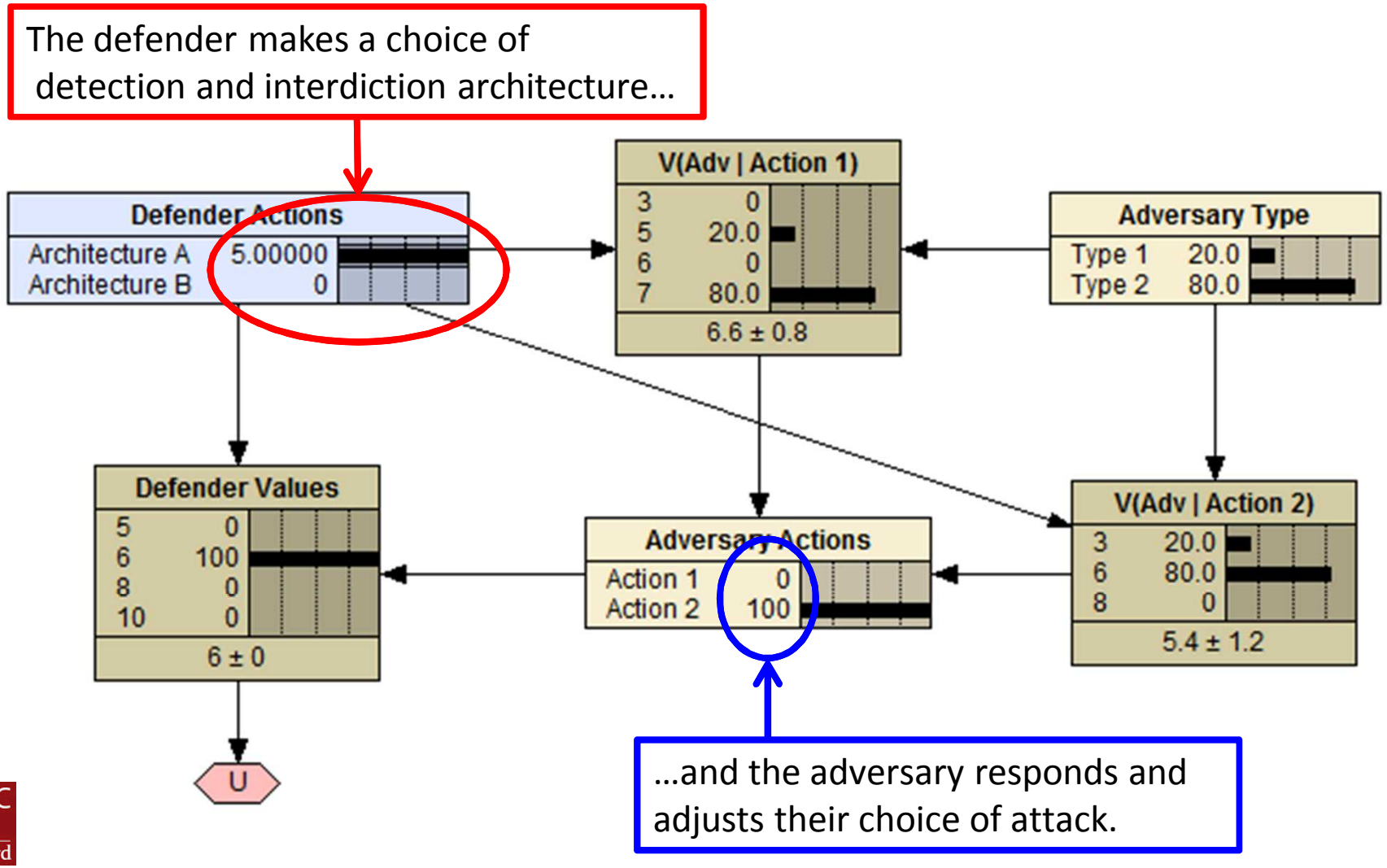
This model can be implemented using familiar tools.

Analysis can then answer important questions.



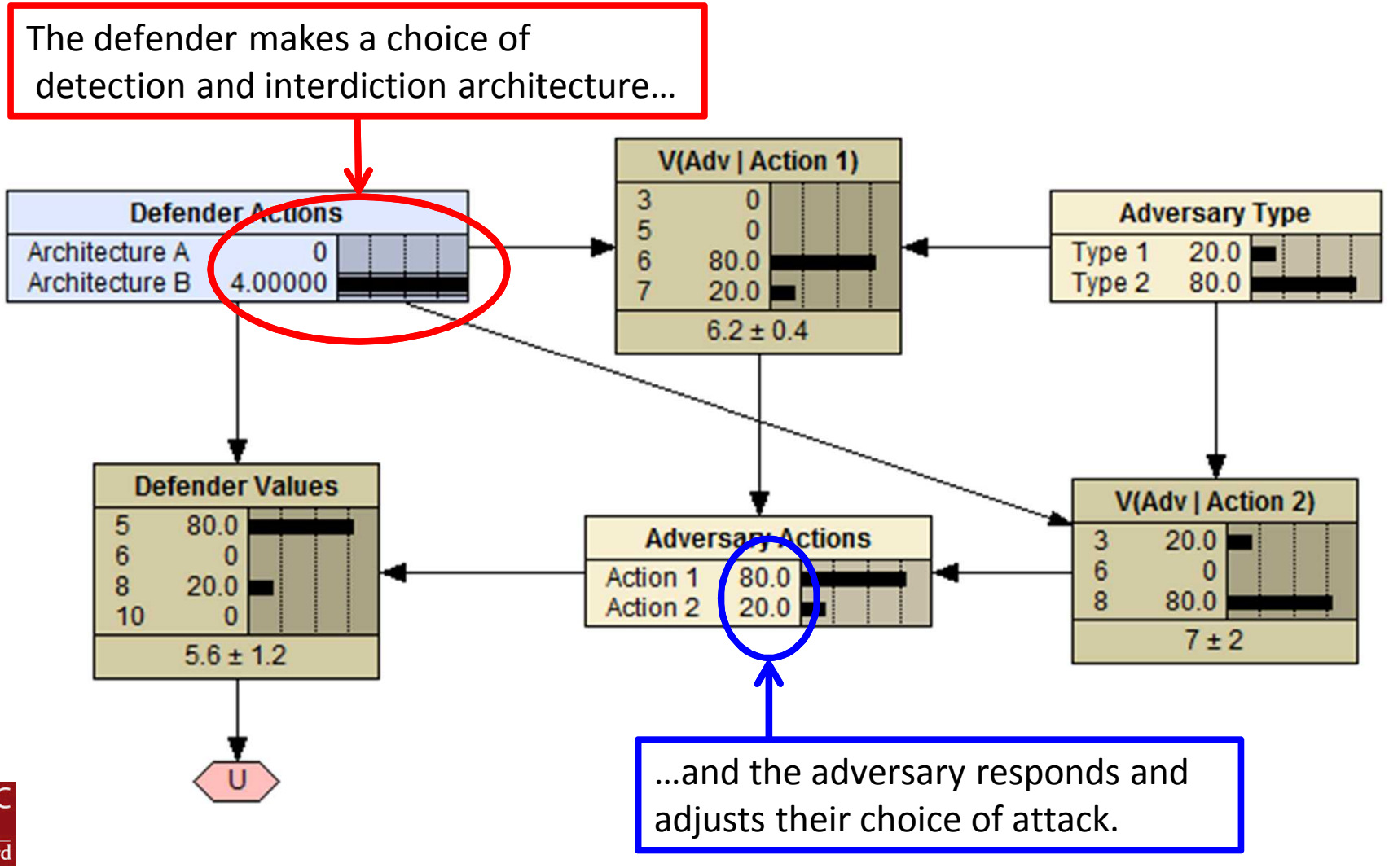
CISAC Stanford

Threat shifting behavior is captured by the game theoretic model.



CISAC Stanford

Threat shifting behavior is captured by the game theoretic model.



CISAC Stanford

Game Theoretic Methods: Assessment

- Assumptions
  - Adversaries make decisions according to a definable set of rules (e.g. utility maximization)
  - Defender can express their beliefs about adversary preferences and outcomes
- Pros
  - Allows for adaptive adversary behavior, including “deterrence”
  - Formalizes information and intuition about adversary behavior
  - Can be extended to allow for increasing adversary sophistication (learning, repeated games, etc.)
- Cons
  - Can be complicated to solve, and computationally expensive
  - Game theory experts can often focus on equilibrium solutions that make unreasonable assumptions

Probability or Game Based Methods: Selected References

1. Jason Merrick and Gregory Parnell, "A Comparative Analysis of PRA and Intelligent Adversary Methods for Counterterrorism Risk Management," Risk Analysis, Vol. 31, No. 9, 2011.
2. Rios Insua D, Jesus J, Banks D, "Adversarial Risk Analysis," Journal of the American Statistical Association," Vol. 104, No. 486, 2009
3. Robert Aumann, "Correlated Equilibrium as an Expression of Bayesian Rationality," Econometrica, Vol. 55, No. 1, Jan., 1987, pp. 1-18
4. Jules van Binbsbergen and Leslie Marx, "Exploring Relations Between Decision Analysis and Game Theory," Decision Analysis, Vol. 4, No. 1, March 2007, pp. 32–40

For more advanced methods, see, for example:

1. Xia Qu and Prashant Doshi, Individual Planning in Infinite-Horizon Multi-agent Settings: Inference, Structure and Scalability, in NIPS 2015

Conclusions

- Modeling adversaries allows for more complete assessments of risk to inform decision making.
- There are many reasonable ways to model adversary decisions, each with pros and cons, but no “correct” ways.
- Red Team and Difficulty Based methods can help uncover vulnerabilities and unexpected outcomes, but lack mathematical formality.
- Risk and Game Theoretic methods formalize adversary decisions mathematically, but can not “predict” adversary actions.
- Most importantly, modeling adversaries allows us to test assumptions about adversary behavior and understand how risk depends on adversary actions.

Image Credits and References

- https://commons.wikimedia.org/wiki/File:FEMA_-_33203_-_Top_Off-4_Exercise_workers_in_Seattle,_WA.jpg
- https://www.llnl.gov/news/maritime-exercise-shows-radiation-detection
- http://www.amazon.com/Correlates-War-v-1/dp/0029289602
- http://www.ucpress.edu/book.php?isbn=9780520247093