Accountability and Identifiability

Joan Feigenbaum,[1] Aaron D. Jaggard,[2 (now 3)] and Rebecca N. Wright[2]

[1]Yale [2]Rutgers [3]NRL http://dimacs.rutgers.edu/~adj/accountability/

NSF awards CNS-1016875 and CNS-1018557

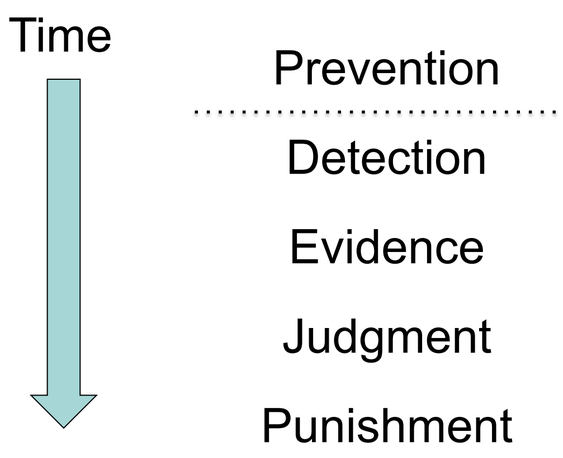
- Accountability complements notions of preventive security
- Agreement that "accountability" is a good thing, but disagreement about what it is
  - Various approaches in CS, administrative law, and international relations
- Formalize accountability so that we can start to reason about it
- Goals: Clarify accountability and understand its relationship to identity

Working Definition

An entity is *accountable* with respect to some policy if, whenever the entity violates the policy, then it is, or could be, punished (perhaps probabilistically)

This intentionally avoids requiring external action (*e.g.*, "hold … responsible"); we want to allow for passive ("automatic") punishment.

This shifts the focus to punishment, moving away from earlier (post-violation) aspects that are more closely tied to identity.



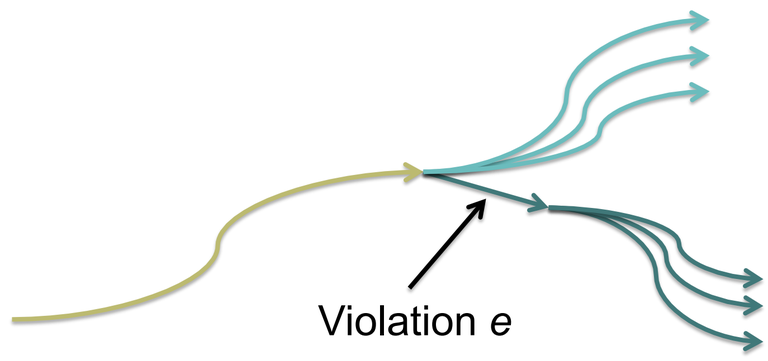
Formalizing Accountability

- System behavior as event traces
- Utility functions for participants
  - Might only know distribution of utilities or have a measure of how "typical" a utility function is
- Principal(s) associated to events
- What qualifies as "punishment" for a violation?
- Punishment should ignore "luck" and be related to the violation in question

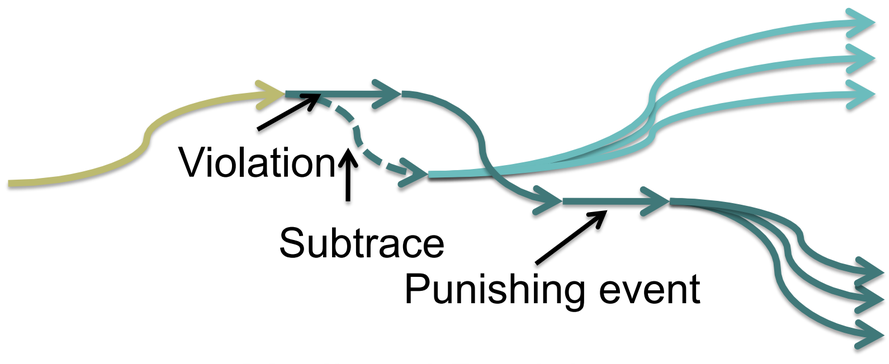
Automatic and Mediated Punishment

- Related to violation, but may or may not involve a punishing action
- Automatic: Is the violator's utility lowered (in expectation or w.r.t. a "typical" utility) after committing the violation?
  - This might reduce the need for identifiability!
- Mediated: Punishing event causally related to the fact of the violation
  - Compare effect of this on utility with outcomes that ignore the violation and things causally dependent upon it

Deterrence and what it means for punishment to be "effective"



Automatic Punishment



Mediated Punishment

Interested in meeting the PIs? Attach post-it note below!