CSC 495.002 – Lecture 9
AI for Privacy: Privacy Norms

Dr. Özgür Kafalı

North Carolina State University
Department of Computer Science

Fall 2017

Agents and Reasoning

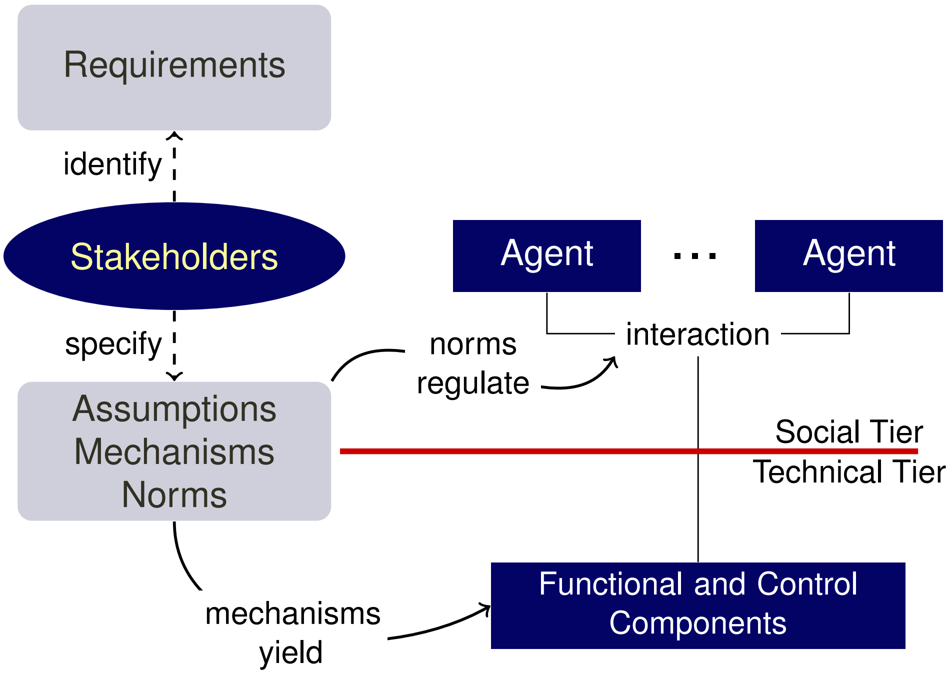
- Agents in pervasive healthcare
- Resolving multi-party privacy concerns via argumentation
- Negotiating privacy preferences

Problem Definition

- Imagine you are developing a healthcare application
- You designed a perfect role-based access control mechanism to regulate access to sensitive patient information
- But, you later observed nurses are sharing their passwords to access each other's accounts

- You cannot control everything with software features
- Provide flexibility to users (don't prevent everything)
- You need a social mechanism to regulate the interactions among users
- Hold users accountable for their actions

Sociotechnical Systems (STS)



Requirements

identify

Stakeholders

specify

Assumptions
Mechanisms
Norms

Agent

...

Agent

interaction

norms
regulate

Social Tier

Technical Tier

mechanisms
yield

Functional and Control Components

Objectives

- Develop abstractions, models, and tools to help address legal and social aspects of security and privacy
- Build computational models of the social architecture
- Enable unified treatment of technical and social considerations

STS Example: Hospital Organization

- Roles: Physician, hospital, patient
- Assumptions: Physicians cannot authenticate when there is a power failure
- Mechanisms: Hospital software allows physicians to authenticate with valid passwords
- Norms: Physicians should not disclose patient information to outsiders

Exercise: Course Management System

- Roles?
- Assumptions?
- Mechanisms?
- Norms?

Contextual Integrity

- A conceptual framework to evaluate the flow of information between parties
- Norms change depending on context
- Previous example: Physicians should not disclose patient information to outsiders
- Are there any variations of this norm? If the context changes
- Physicians may disclose patient information to family members in emergencies

Barth et al. Privacy and Contextual Integrity: Framework and Applications. IEEE Symposium on Security and Privacy, pages 184–198, 2006

Formal Specification

- N(SUBJECT, OBJECT, antecedent, consequent)

- Type: $N \in \{$Commitment (C), Authorization (A), Prohibition (P)$\}$
- SUBJECT: Party that is [responsible for / beneficiary of] the norm
- OBJECT: Party that is [beneficiary of / responsible for] the norm
- antecedent: Preconditions that need to hold to activate the norm
- consequent: Action that needs to be [performed / avoided]

Commitment

- Informally, describes “what you should do”

- Example: A physician is committed to the hospital to operating upon patients in an emergency

- Formally, $\boxed{C(\text{PHYSICIAN}, \text{HOSPITAL}, \text{emergency}, \text{operate})}$

Authorization

- Informally, describes “what you can do”
- Example: A physician is authorized by the hospital to access the patient’s electronic health records (EHR) if the patient gives consent
- Formally, $\boxed{A(\textsc{Physician}, \textsc{Patient}, \text{consent}, \text{view\_EHR})}$

Prohibition

- Informally, describes “what you should not do”
- Example: A physician is prohibited by the hospital from disclosing a patient’s protected health information (PHI) to others
- Formally, $\boxed{P(\textsc{Physician}, \textsc{Hospital}, \text{true}, \text{disclose\_PHI})}$

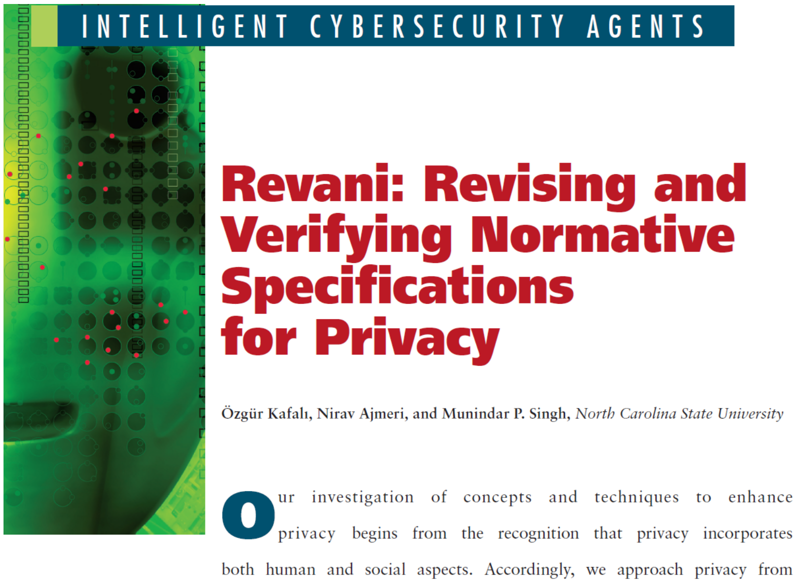
Exercise: Norm Specifications

- A physician may prescribe drugs to the patients or schedule their next visit after a routine visit
  $\boxed{A(\text{PHYSICIAN}, \text{HOSPITAL}, \text{visit}, \text{prescribe} \lor \text{schedule\_visit})}$

- Hospital workers must log out of a public computer as soon as they finish viewing EHR of patients
  $\boxed{C(\text{WORKER}, \text{HOSPITAL}, \text{public\_computer} \land \text{view\_EHR}, \text{logout})}$

- A nurse should not disclose patient information to patient's family unless there is consent from the patient or it's an emergency
  $\boxed{P(\text{NURSE}, \text{HOSPITAL}, \neg\text{consent} \land \neg\text{emergency}, \text{disclose\_family})}$

Normative Specifications for Privacy



INTELLIGENT CYBERSECURITY AGENTS

Revani: Revising and Verifying Normative Specifications for Privacy

Özgür Kafalı, Nirav Ajmeri, and Munindar P. Singh, *North Carolina State University*

Our investigation of concepts and techniques to enhance privacy begins from the recognition that privacy incorporates both human and social aspects. Accordingly, we approach privacy from

Kafalı et al. Revani: Revising and Verifying Normative Specifications for Privacy. IEEE Intelligent Systems, 31(5):8–15, 2016

Research Questions

- Specification: What are the necessary components to develop a computational model of an STS?
- Verification: How can we verify that an STS satisfies the functional, security (and privacy) requirements of its stakeholders?
- Refinement: Supposing an STS fails to satisfy its requirements, how can we propose refinement so that its refined specification satisfies the requirements?

STS Components: Assumptions

- Example: Physicians cannot authenticate when there is a power failure
- Formally, $\langle \neg$authenticate, power_failure$\rangle$
  or,
  $\neg$authenticate $\leftarrow$ power_failure

STS Components: Mechanisms

- Example: Hospital software allows physicians to authenticate with valid passwords

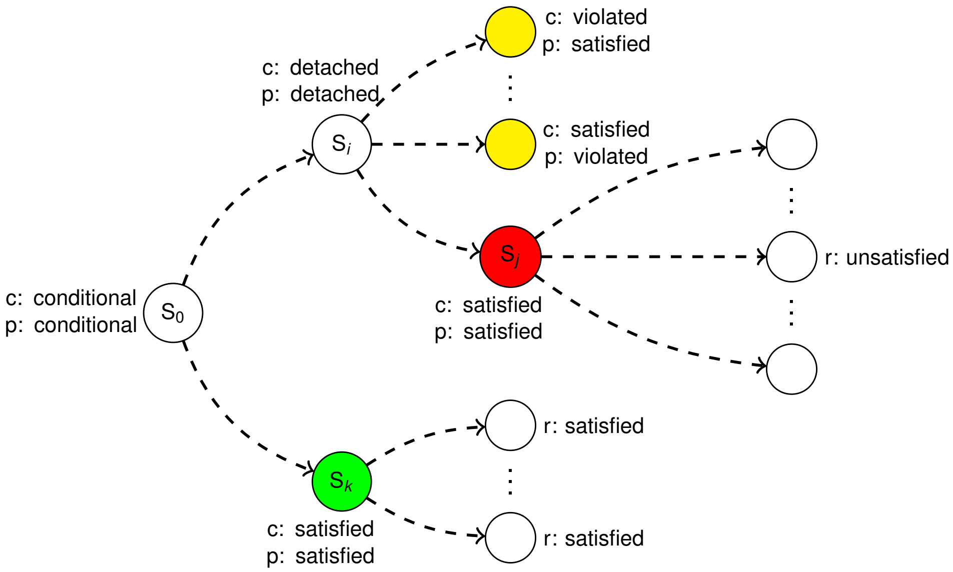
- Formally, m(enabler, add, delete)
- m(enter_password, $\{$authenticate$\}$, $\{\ \}$)

Requirements in Temporal Logic

- Express stakeholder requirements as Computation Tree Logic (CTL) formulas
  - A branch quantifier, all (A) or exists (E), over branches emanating from the current point
  - A linear temporal operator, describing properties of a single branch (next (X), eventually (F), always (G), and until (U))

- Examples:
  - Physicians should always be able to access patients' EHR
    In CTL: $\boxed{\text{AF view\_EHR}}$
  - Physicians should never disclose patients' PHI
    In CTL: $\boxed{\text{AG } \neg\text{disclose\_PHI}}$

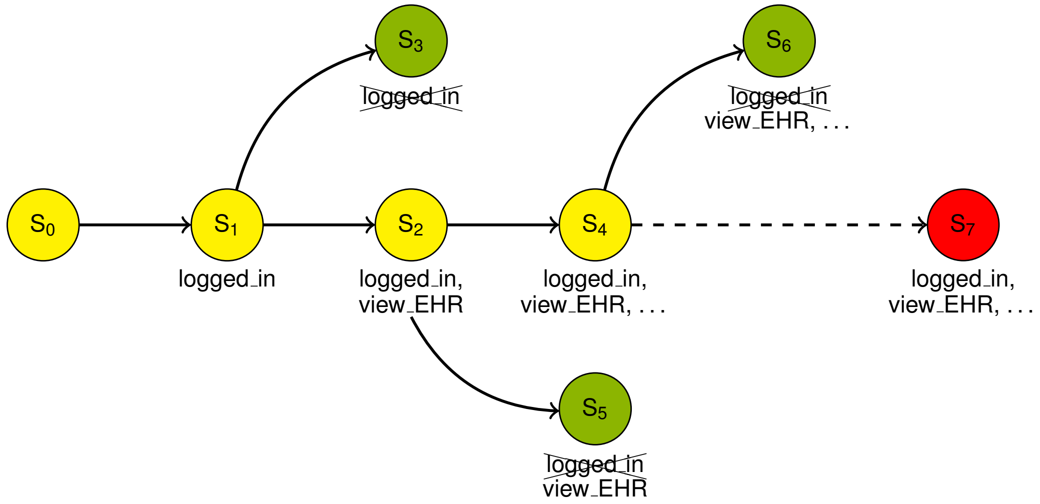
Verification Setting



Verification Example

- Open sessions must be closed after reviewing EHR

$$\boxed{\text{AG (view\_EHR} \rightarrow \text{AF } \neg\text{logged\_in)}}$$



Refinement

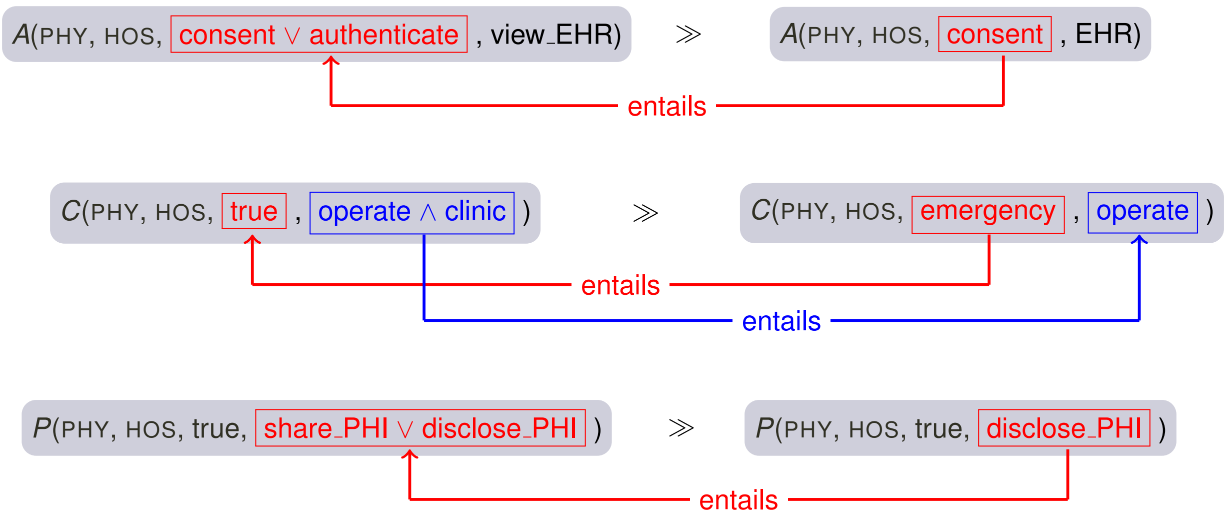
- *Refinement* of a norm: Generalization or specialization of its antecedent or consequent
- An iterative design process to refine norms of an STS specification
  - Takes as input a set of (unsatisfied) requirements
  - Each refinement is captured with a design pattern

Refinement Patterns

- Pattern: A general reusable solution to a commonly occurring problem
- Strengthening: Specify more strict norms
- Weakening: Relax norms
- Amendment: Combine strengthening and weakening
- Overseer: Assign a monitor to a given norm
- Operational: Refine mechanisms
- Sociotechnical: Transform between tiers

https://en.wikipedia.org/wiki/Software_design_pattern

Norm Strength



$A(\text{PHY}, \text{HOS}, \text{consent} \vee \text{authenticate}, \text{view\_EHR}) \gg A(\text{PHY}, \text{HOS}, \text{consent}, \text{EHR})$

entails

$C(\text{PHY}, \text{HOS}, \text{true}, \text{operate} \wedge \text{clinic}) \gg C(\text{PHY}, \text{HOS}, \text{emergency}, \text{operate})$

entails

entails

$P(\text{PHY}, \text{HOS}, \text{true}, \text{share\_PHI} \vee \text{disclose\_PHI}) \gg P(\text{PHY}, \text{HOS}, \text{true}, \text{disclose\_PHI})$

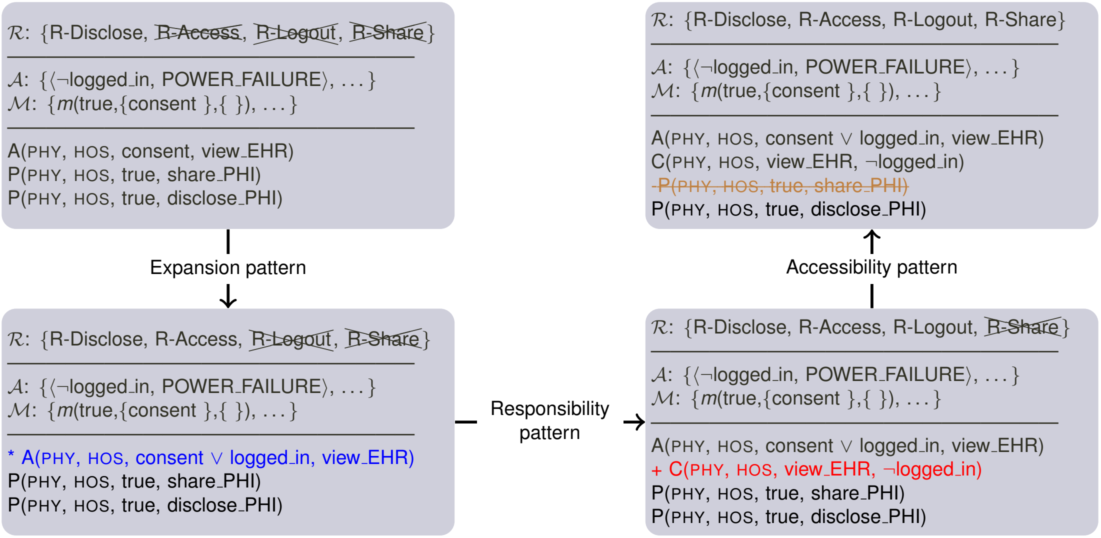
entails

Sample Pattern

- Transform specifications between technical and social tiers
- Relaxing a mechanism may introduce security and privacy risks
- Specify a complementary commitment to mitigate security and privacy concerns
    - Physician is authorized to use PC for 15 minutes before session expires
    - Extend authorization's duration to two hours (technical tier)
    - Physician commits to logging off from computer (social tier)
    - Physician is accountable if commitment violated

Application of Patterns

R-Disclose: AG (¬ disclose_PHI)
R-Access: EF (view_EHR)
R-Logout: AG (view_EHR → AF ¬logged_in)
R-Share: AG (disaster → EF share_PHI)



$\mathcal{R}$: {R-Disclose, ~~R-Access~~, ~~R-Logout~~, ~~R-Share~~}
$\mathcal{A}$: {⟨¬logged_in, POWER_FAILURE⟩, . . . }
$\mathcal{M}$: {*m*(true,{consent },{ }), . . . }
A(PHY, HOS, consent, view_EHR)
P(PHY, HOS, true, share_PHI)
P(PHY, HOS, true, disclose_PHI)

Expansion pattern

$\mathcal{R}$: {R-Disclose, R-Access, ~~R-Logout~~, ~~R-Share~~}
$\mathcal{A}$: {⟨¬logged_in, POWER_FAILURE⟩, . . . }
$\mathcal{M}$: {*m*(true,{consent },{ }), . . . }
* A(PHY, HOS, consent ∨ logged_in, view_EHR)
P(PHY, HOS, true, share_PHI)
P(PHY, HOS, true, disclose_PHI)

Responsibility pattern

$\mathcal{R}$: {R-Disclose, R-Access, R-Logout, ~~R-Share~~}
$\mathcal{A}$: {⟨¬logged_in, POWER_FAILURE⟩, . . . }
$\mathcal{M}$: {*m*(true,{consent },{ }), . . . }
A(PHY, HOS, consent ∨ logged_in, view_EHR)
+ C(PHY, HOS, view_EHR, ¬logged_in)
P(PHY, HOS, true, share_PHI)
P(PHY, HOS, true, disclose_PHI)

Accessibility pattern

$\mathcal{R}$: {R-Disclose, R-Access, R-Logout, R-Share}
$\mathcal{A}$: {⟨¬logged_in, POWER_FAILURE⟩, . . . }
$\mathcal{M}$: {*m*(true,{consent },{ }), . . . }
A(PHY, HOS, consent ∨ logged_in, view_EHR)
C(PHY, HOS, view_EHR, ¬logged_in)
~~P(PHY, HOS, true, share_PHI)~~
P(PHY, HOS, true, disclose_PHI)

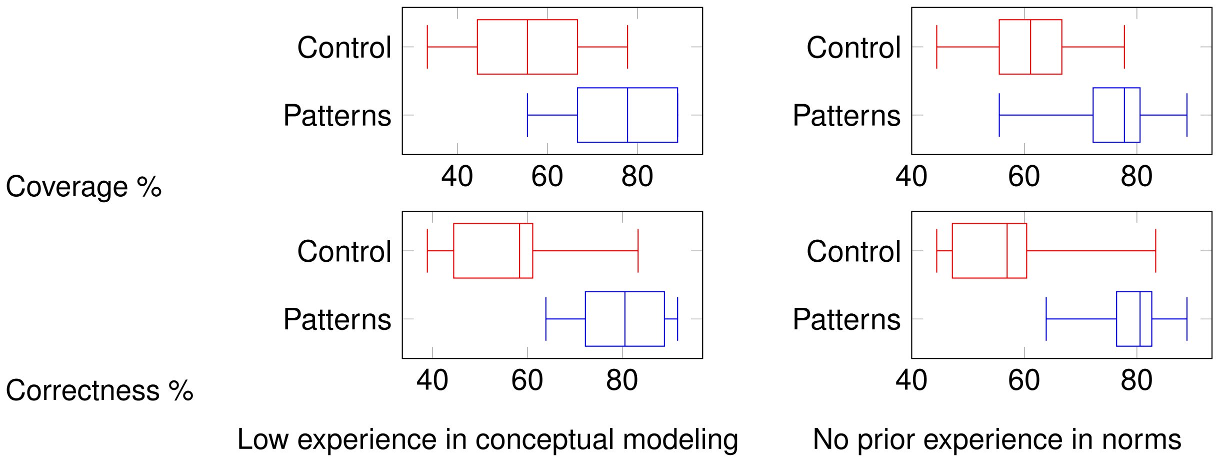
How Much and When do Patterns Help?

- Questions
  - Do patterns help design better STSs given the requirements?
  - Does prior industry experience or knowledge of norms affect quality of design?
- Preliminary study with 32 participants (computer science graduate students)
  - Control group (no patterns) vs treatment group (patterns), balanced in education and experience
  - After a learning phase, each group designs and refines an STS via norms
  - Small scenario; correct solution established by two of the authors

Metrics

- **Coverage** of design: Fraction of norms in the oracle that are stated by the participants in each phase
- **Correctness** of design: Fraction of participant-stated norms that occur in the oracle for each phase
- **Time** to design: Time in minutes recorded by participants to complete each phase
- **Ease** of design: Subjective ratings provided by the participants via a post-study survey (Likert scale, 1–5)

Results


Control
Patterns
Control
Patterns
Coverage %
Control
Patterns
Control
Patterns
Correctness %
Low experience in conceptual modeling
No prior experience in norms