A Logic of Privacy

Anupam Datta
Carnegie Mellon University

Formalizing Privacy Policies governing Individual Information

What is Privacy?

Contextual Integrity

Formalize and Enforce Privacy Policies

Individual Information

A Logic of Privacy

Aggregate Statistics

Cryptography, Statistics

Privacy Laws in the US

- HIPAA (Health Insurance Portability and Accountability Act, 1996)
  - Protecting personal health information
- GLBA (Gramm-Leach-Bliley-Act, 1999)
  - Protecting personal information held by financial service institutions
- COPPA (Children‘s Online Privacy Protection Act, 1998)
  - Protecting information posted online by children under 13

Privacy Policies

- Most major companies have privacy policies

Representing Complex Privacy Laws

Challenges

- Identifying core privacy concepts in long, dense legal text
  - HIPAA has 84 operational clauses about disclosures of protected health information (~30 pages)
- Understanding how individual clauses should be combined
  - permitting clauses, denying clauses, cross-references, exceptions

Main Result

1. PrivacyLFP, a first-order logic (language) for representing privacy laws

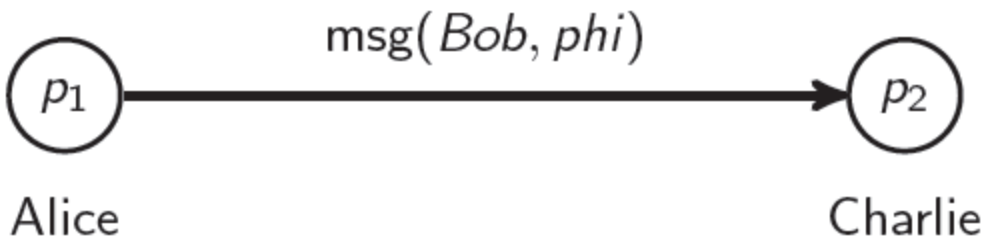
2. First complete logical formalization of all disclosure-related clauses in the HIPAA Privacy Rule and the Gramm-Leach-Bliley Act

H. DeYoung, D. Garg, L. Jia, D. Kaynar, A. Datta, Experiences in the Logical Specification of the HIPAA and GLBA Privacy Laws, in *WPES*, October 2010.

Builds on:

A. Barth, A. Datta, J. C. Mitchell, H. Nissenbaum, Privacy and Contextual Integrity: Framework and Applications, in *S&P* , May 2006.

Transmission of Personal Information



Transmissions of this form are governed by contextual norms (Nissenbaum 2004)

Example from HIPAA Privacy Rule

A covered entity may disclose an individual's protected health information (phi) to law-enforcement officials for the purpose of identifying an individual if the individual made a statement admitting participating in a violent crime that the covered entity believes may have caused serious physical harm to the victim

- Concepts in privacy policies
  - **Actions:** send(p1, p2, m)
  - **Roles:** inrole(p2, law-enforcement)
  - **Data attributes:** attr_in(prescription, phi)
  - **Temporal constraints:** in-the-past(state(q, m))
  - **Purposes:** purp_in(u, id-criminal))
  - **Beliefs:** believes-crime-caused-serious-harm(p, q, m)

Black-and-white concepts

Grey concepts

Example HIPAA Clause

A covered entity may disclose an individual's protected health information (phi) to law-enforcement officials for the purpose of identifying an individual if the individual made a statement admitting participating in a violent crime that the covered entity believes may have caused serious physical harm to the victim

$$
\begin{aligned}
&\forall p1, p2, m, u, q, t. \\
&\quad (\text{send}(p1, p2, m) \wedge \\
&\quad\ \text{inrole}(p2, \text{law-enforcement}) \wedge \\
&\quad\ \text{tagged}(m, q, t, u) \wedge \\
&\quad\ \text{attr_in}(t, \text{phi})) \\
&\quad\ \supset (\text{purp_in}(u, \text{id-criminal})) \\
&\qquad \wedge \exists\, m'. \Diamond\text{state}(q,m') \wedge \text{is-admission-of-crime}(m') \\
&\qquad \wedge \text{believes-crime-caused-serious-harm}(p1, q, m')
\end{aligned}
$$

Combining Clauses

- Two types of clauses
  - Positive norm: disclosure permitted *if* requirement satisfied
    - "A covered entity may disclose protected health information for treatment activities […]" [HIPAA 164.506(c)(2)]
  - Negative norm: disclosure permitted *only if* requirement satisfied
    - "A covered entity must obtain authorization for any use or disclosure of psychotherapy notes." [HIPAA 164.508(a)(2)

- A disclosure is permitted if it satisfies a*t least one* positive norm and *all* the negative norms

$$maysend(p_1, p_2, m) \triangleq \left(\bigvee_i \varphi_i^+\right) \wedge \left(\bigwedge_j \varphi_j^-\right)$$

Exceptions to norms

Exceptions to negative norms:
"A covered entity must obtain an authorization for any use or disclosure of psychotherapy notes, except [...]."

Conclusion: Satisfy either the core <u>or</u> one of the exceptions.

$$\varphi^{-}_{\texttt{164.508a2}'} \triangleq \varphi^{-}_{\texttt{164.508a2}} \vee (\varphi^{e}_{\texttt{164.508a2iA}} \vee \cdots)$$

"Exceptions" to positive norms:

- A covered entity may disclose information to report abuse.
- Disclosures under previous require informing the victim.

Conclusion: Satisfy the core <u>and</u> its refinements.

$$\varphi^{+}_{\texttt{164.512c1}'} \triangleq \varphi^{+}_{\texttt{164.512c1}} \wedge \varphi^{e}_{\texttt{164.512c2}}$$

Structure of HIPAA and GLBA

- HIPAA Privacy Rule
  - Deny all transmissions not explicitly allowed
  - 56 positive norms, 7 negative norms, 19 exceptions
  - Formalization in logic: 94 pages with explanation
- GLBA
  - Allow all transmissions not explicitly denied
  - 5 negative norms and 10 exceptions
  - Formalization in logic: 12 pages with explanation

- Important property of formalization
  - Traceability: Each clause in law corresponds to one norm or exception in formalization (roughly)

Syntax of Policy Logic

$$
\begin{array}{llll}
\text{Atoms} & P & ::= & p(t_1, \ldots, t_n) \\
\text{Formulas} & \varphi & ::= & P \mid \top \mid \bot \mid \\
 & & & \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \\
 & & & \forall \vec{x}.(c \supset \varphi) \mid \exists \vec{x}.(c \wedge \varphi) \\
\text{Restrictions} & c & ::= & P \mid \top \mid \bot \mid c_1 \wedge c_2 \mid \\
 & & & c_1 \vee c_2 \mid \exists x.c
\end{array}
$$

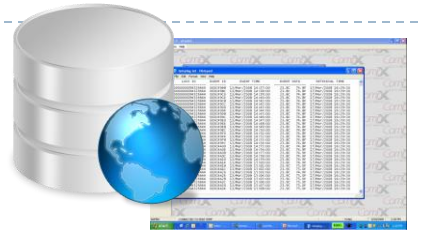
- First-order logic with restricted quantification
- Can express temporal properties, subjective predicates

Related Work: Privacy Policy Specification

- Logics and languages for specification of privacy policies
  - P3P [Cranor et al.], XACML [OASIS], EPAL [Backes et al.], requirements engineering [Breaux and Antón], LPU [Barth et al.], Privacy APIs [Gunter et al.], deontic logic [I. Lee et al.], SecPAL [Becker et al.], ...
- Formal specification of privacy laws
  - LPU [Barth et al.]: Examples from HIPAA and GLBA
  - Datalog HIPAA [Lam et al.]: HIPAA §§164.502, 506, and 510
  - Privacy APIs [Gunter et al.]: HIPAA §164.506
  - Deontic logic [I. Lee et al.]: Examples from FDA CFR §610.40

Formalizing and Enforcing Privacy Policies governing Individual Information

Privacy Policy

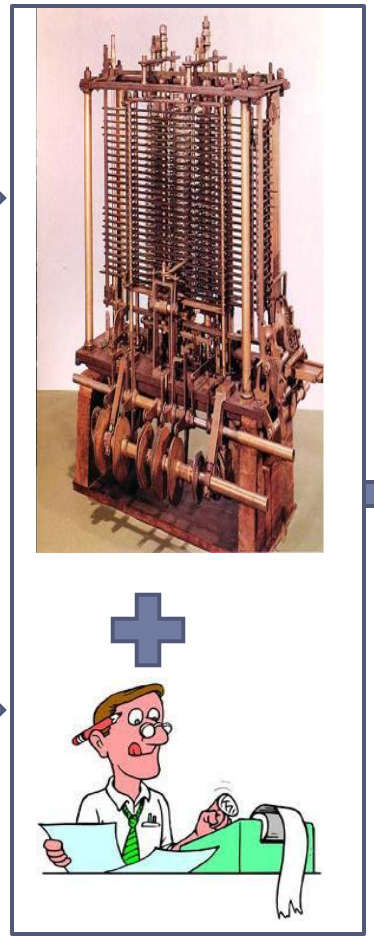


Organizational audit log

Complete formalization of HIPAA, GLBA

$$\mathbf{G}\,\Big(\forall p_1, p_2, m.\ \mathsf{send}(p_1, p_2, m) \supset (\forall d, u, q, t.\ (m = \mathrm{info}(d, u)) \wedge \mathsf{contains}(m, q, t) \supset (\bigvee_i \varphi_i^+) \wedge (\bigwedge_j \varphi_j^-)) \wedge (\forall t.\ (m = \mathrm{req_for_access}(p_1, t)) \supset \varphi^-_{164.524b2i'} \wedge \varphi^-_{164.524b2ii'})\Big)$$

Computer-readable privacy policy



Audit

Automated audit for black-and-white policy concepts

Detect policy violations

Learning to audit for grey policy concepts

Thanks!
Questions?