Deriving Semantic Models from Privacy Policies

Travis D. Breaux and Annie I. Antón

North Carolina State University
{tdbreaux, aianton}@eos.ncsu.edu

POLICY 2005, June 6th 2005

Presentation Outline

- Research Motivation
  - Machine-enforceable policies that comply with law
- Overview, express:
  - Policies as Goals
  - Goals as Restricted Natural Language Statements (RNLS)
  - RNLS(s) as Semantic Models
- Research Results
  - Example Semantic Models
  - Queries over Top 100 Goals
- Current and Future Work
- Research Summary

Towards Machine-enforceable Policies

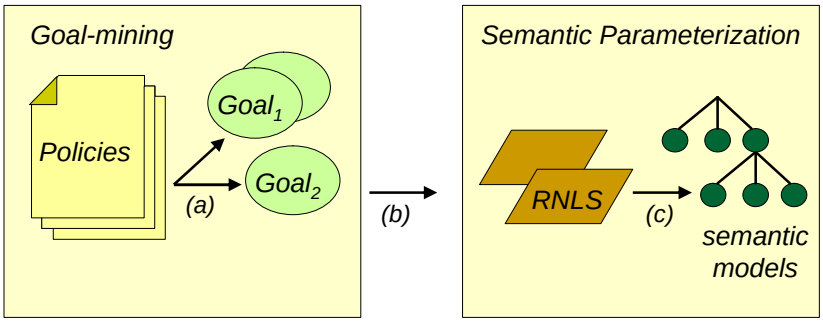
- Motivations
  - Privacy laws require companies to enforce their policies.
  - Consumers are increasingly concerned about privacy violations.
  - Companies are increasingly being held accountable for their privacy practices.

...without machine-readable and machine-enforceable policies, privacy practices will continue to be inconsistently applied and therefore prone to violations

Need a Policy Language that can…

- Represent Rights and Obligations.
  - Rights, like permissions, describe what people and systems ***are allowed to*** do.
  - Obligations describe what people and systems ***must*** do.
- Interface to natural language, policies must...
  - Be maintainable by non-technical policy analysts.
  - Be implementable by system administrators.
  - Be legally enforceable by a court of law.
- Interface to program execution, policies must...
  - Exclusively decide policy-governed control flow.
  - Associate governance semantics with data.

From Policies to Semantic Models



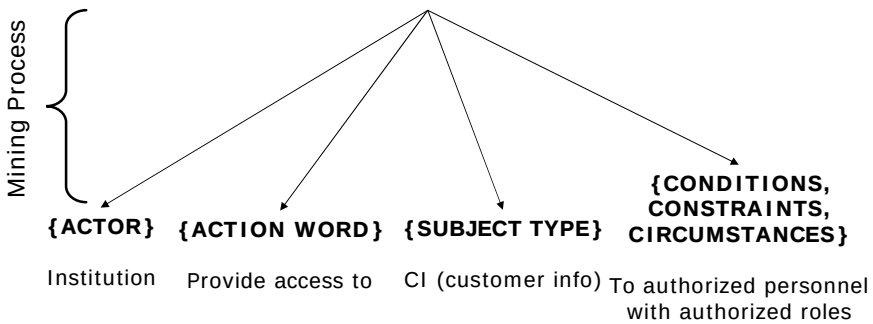
(a) Goals are mined from policies.

(b) Restate goals as Restricted Natural Language Statements (RNLS).

(c) RNLS are parameterized to build semantic models.

Representing Privacy Goals

Privacy Statement: Employees are authorized to access customer information only when they need it, to provide you with accounts and services or to maintain your accounts.



Identify goals using action keywords …

The meaning and use of action keywords in goals is strictly controlled to remove ambiguity.

ACCESS	CONNECT	DISCLOSE	MAINTAIN	INVESTIGATE	RESERVE
AGGREGATE	CONSOLIDATE	DISPLAY	MAKE	POST	REVIEW
ALLOW	CONTACT	ENFORCE	MAXIMIZE	PREVENT	SHARE
APPLY	CONTRACT	ENSURE	MINIMIZE	PROHIBIT	SPECIFY
AVOID	CUSTOMIZE	EXCHANGE	MONITOR	PROTECT	STORE
BLOCK	DENY	HELP	NOTIFY	PROVIDE	UPDATE
CHANGE	DESTROY	HONOR	OBLIGATE	RECOMMEND	URGE
CHOOSE	DISALLOW	IMPLY	OPT-IN	REQUEST	USE
COLLECT	DISCIPLINE	INFORM	OPT-OUT	REQUIRE	VERIFY
COMPLY	DISCLAIM	LIMIT			

Source: *Privacy Goal Management Tool, NCSU, IEEE Security & Privacy,* 2004

From Goals to Restricted Natural Language Statements (RNLSs).

- The full scope of natural language is too complex!
- Each RNLS describes one activity with external references to other RNLSs.
- Rights and obligations are described by activities.

Goal: (Provider, SHARE information to market services.)

RNLS #1: The provider markets services.
RNLS #2: The provider may share information to (RNLS#1).

Our Semantic Models

- For our purposes, semantic models are…
  - Structural representations of meaning.
  - Sufficiently unique to differentiate concepts.
  - Amenable to asking *what*, *when*, *why* and *how* questions.
- Models are built from three formal relations:
  - σ - unary, *root relation* (main idea or concept).
  - α - binary, *associative relation* (conceptual relations).
  - δ - binary, *declarative relation* (values assigned to conceptual relations).

Example Semantic Model

- ***RNLS #3***: The provider may share information with whom?

σ(*activity*)
α(*activity*, *actor*)
α(*activity*, *action*)
α(*activity*, object)
α(*activity*, *target*)
δ(*actor*, *provider*)
δ(*action*, *share*)
δ(*object*, *information*)
δ(*target*, *?whom*)

- The modal "may" indicates a *right*.

α(*provider*, *right*)	δ(*right*, *activity*)

- The semantic model in the CFG:

```
activity [ right : provider ] {
      actor = provider
      action = share
      object = information
      target = ?whom
}
```

Queries Across the Top 100 Goals

ID	Object	Target
155	transaction information	subsidiary
155	experience information	subsidiary
822	PII	affiliate
822	PII	service-provider
954	information	third-party
954	statistics	third-party
156	transaction information	affiliate
156	experience information	affiliate
170	PII	subsidiary

Reflexive Models: Purpose and Instruments

- ***RNLS #5:*** The provider may use cookies to collect information.
- ***RNLS #6:*** The provider may collect information using cookies.

```
activity [ right : provider ] {
    actor = provider
    action = use
    object = cookie
    purpose = activity {
        action = collect
        object = information
    }
}
```

```
activity [ right : provider ] {
    actor = provider
    action = collect
    object = information
    instrument = cookie
}
```

Current and Future Work

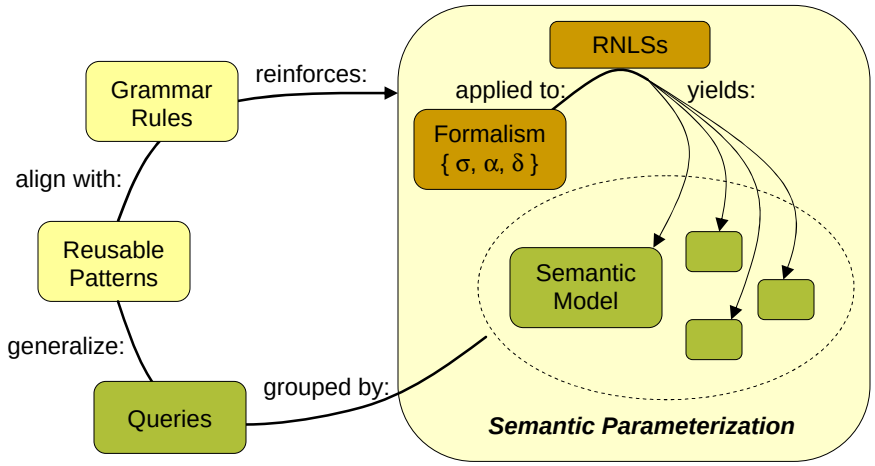
- Apply Semantic Parameterization to law to…
  - Identify rights and obligations.
  - Identify rules for business processes and systems.
- Working with the U.S. Law: Health Insurance Portability and Accountability Act (*HIPAA*).
  - Pilot Study: *The HIPAA Fact Sheet: Protecting the Privacy of Patient's Health Information*
  - Case Study: *The HIPAA Privacy Rule*, enforced by the Department of Health and Human Services.

Future Work: Example Rule

- Providers ***will*** <provide the patient access to their medical records> ***within*** <30 days of the patient's request>.
- Semantic models for two activities as events:
  - M1: Patient requests access (via right).
  - M2: Provider provides access (via obligation).
- Unit of time: 30 days.

Rule: if { M1 } then { M2 $<_{time}$ { 30 days $+_{time}$ M1 } }

Developing a Repeatable, Iterative Process



In Summary…

- Contribution
  - Provides new structure for modeling policy goals.
  - Support for querying policy statements.
- Limitations
  - CFG requires new semantics for representing rules.
  - The subjectivity of the parameterization process must be validated.
- Future Work
  - Empirical studies to validate semantic parameterization.
  - Analysis of law governing information sharing practices.
  - Investigate models to align NL policies with systems.

Feedback and Questions?

Travis D. Breaux and Annie I. Antón

To see more of our work, visit our website:
http://ThePrivacyPlace.org