Analyzing Goal Semantics for Rights, Permissions and Obligations

Travis D. Breaux and Annie I. Antón

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

RE 2005, September 1st 2005

Presentation Outline

- Research Motivation
- Expressing: Goals, RNLSs, Semantic Models
- Example Models and Queries
- Natural Language Generation
- Future Work & 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.

- Problem Statement

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

Relating Policies to Requirements…

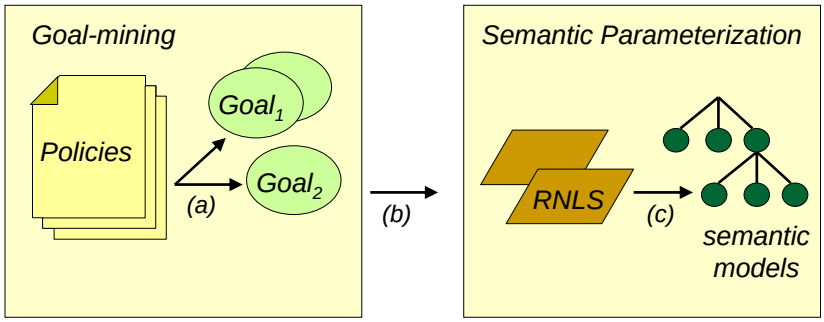
- Policies describe ***both*** requirements of systems ***and*** responsibilities of people.
- Some responsibilities are implementable through system requirements.
- These responsibilities are either machine-enforceable while others are only machine-accountable.

… by modeling both responsibilities and system requirements we seek to describe the whole policy picture that ensures policy-compliant systems.

Need a policy language that can…

- Represent rights and obligations.
  - Rights, like permissions, describe what people and systems ***may*** or ***may not*** 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



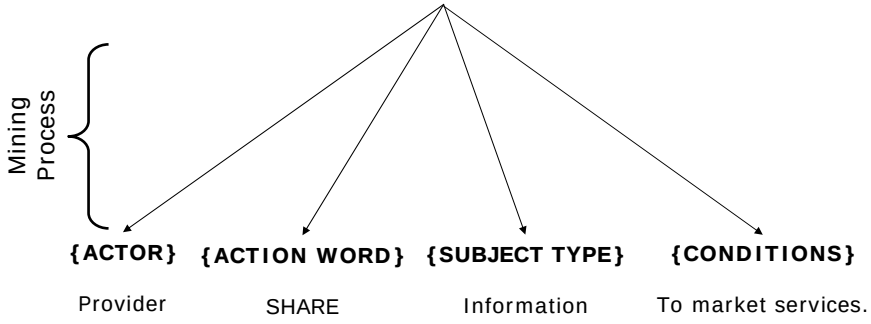
(a) Goals are mined from policies.

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

(c) RNLS are parameterized to build semantic models.

Representing Privacy Policies as Goals

Privacy Statement: We share information for the purpose of marketing services.



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

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 1.1: The provider markets services.
RNLS 1.2: The provider may share information to (RNLS#1).

Our Semantic Models

- For our purposes, semantic models are…
  - Structured representations of meaning.
  - Sufficiently unique to differentiate concepts.
  - Amenable to asking *what*, *when*, *why* and *how* questions.
- Models are a triple $\langle \sigma, A, \Delta \rangle$:
  - σ - unary, *root relation* (main idea or concept)
  - α - binary, *associative relation* (conceptual relation)
  - δ - binary, *declarative relation* (maps conceptual relation to values)

Simple Semantic Model

- ***RNLS 2***: 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*)

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


What benefit do these models provide?

Targeted and Open-ended Queries

- Two types of queries:
  - Boolean queries - pair-wise relational match.
  - *Wh*-queries - pair-wise relational match and variables store corresponding values as query responses.
- Example:
  - ***What*** information may be shared with ***whom***?

ID	***Object***	***Target***
155	transaction information	subsidiary
156	experience information	affiliate
954	statistics	third-party

Reflexive Models: Purpose and Instruments

- ***RNLS 3.1:*** The provider may use cookies to collect information.
- ***RNLS 3.2:*** The provider may collect information using cookies.

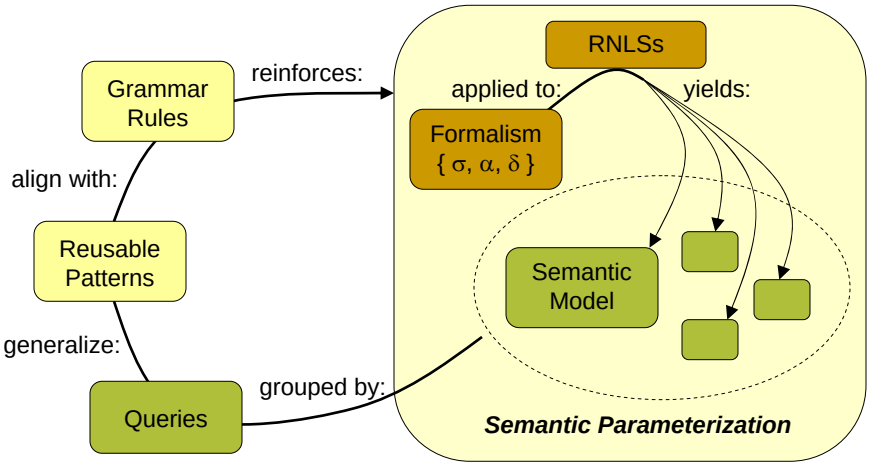
```
activity [ right : provider ] {
    actor = provider
    action = use
    object = cookie
    purpose = activity {
        action = collect
        object = information
    }
}
```

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

Template Method for Natural Language (NL) Generation

- Associates a class of semantic models with a natural language statement.
- Templates are a pair ⟨ *Q*, *S* ⟩:
  - *Q*: a unique Boolean or *Wh*-query.
  - *S*: a parameterized NL statement.
- Example Statement:
  - The ***?subject*** may share ***?object*** with ***?target***.
- Limitations (and future work):
  - Conditional syntax in English requires special sub-queries (e.g., possessive forms, conjunctions, etc.)

Developing a Repeatable, Iterative Process



Current and Future Work

- Applying 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.

Example from HIPAA Privacy Rule
[*WPES'05*]

- 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 } }

In Summary…

- Contributions
  - New structure for modeling policies and requirements.
  - Support for organizing requirements (CFG + Tool).
- Limitations
  - CFG requires semantics for conditions, constraints, etc.
  - The subjectivity of semantic parameterization must be evaluated.
- Future Work
  - Empirical studies to validate semantic parameterization.
  - Analysis of law governing information sharing practices.
  - Investigate models to align policies with systems.

Feedback and Questions?

To see more of our work, visit our website:

http://ThePrivacyPlace.org