Research Directions for Policies and Compliance in Financial Services: Leveraging Semantic Web and Web Services

Benjamin Grosof
Douglas Drane Assistant Professor of Information Technology
MIT Sloan School of Management
http://ebusiness.mit.edu/bgrosof

including joint work with Stuart Madnick, Aykut Firat, Chitravanu Neogy, Said Tabet

Presentation at the Center for eBusiness @ MIT Annual Sponsors Conference
Outline

• Introduction
  – Challenge of Semantics
  – Opportunities of the New Generation Web

• Policies and Compliance
  – Landscape Today
  – Advantages of Standardized Semantic Web Rules
    (Situated Courteous Logic Programs in RuleML)

• Financial Information Integration
  – Conflicting Definitions in Business Reporting
  – Mapping Approach (Extended COntext INterchange)

• Research Challenges & Directions
Challenge: Capturing Semantics

- Deep challenge is to capture the semantics of data and processes, so that can:
  - Represent, monitor, and enforce policies – e.g., trust and contracts
  - Map between definitions of entities, e.g., in reporting
  - Integrate information powerfully
Opportunity from Semantic Web Services -- the New Generation Web Platform

• New technologies for Rules (RuleML standard, based on Situated Courteous Description Logic Programs knowledge representation)
  – + New technologies for Ontologies* (OWL standard)
  – + Databases (SQL, XQuery, RDF)
  – + Web Services (WSDL, SOAP, J2EE, .Net)

• Status today:
  – Technologies: emerging, strong research theory underneath
  – Standards activities: intense (W3C, Oasis, …)
  – Commercialization: early-phase (majors in alpha, startups)

(* Ontology = structured vocabulary, e.g., with subclass-superclass, domain, range, datatypes. E.g., database schemas.)
Next Generation Web

Semantic Web Services

Semantic Web techniques
- Automated Knowledge Bases
- Rules (RuleML)
- Ontologies (OWL)
- Databases (SQL, XQuery, RDF)

Web Services techniques
- API’s on Web (WSDL, SOAP)

XML

First Generation Web

Two interwoven aspects:
- Program: Web Services
- Data: Semantic Web
Outline

• Introduction
  – Challenge of Semantics
  – Opportunities of the New Generation Web

• Policies and Compliance
  – Landscape Today
  – Advantages of Standardized Semantic Web Rules (Situated Courteous Logic Programs in RuleML)

• Financial Information Integration
  – Conflicting Definitions in Business Reporting
  – Mapping Approach (Extended COntext INterchange)

• Research Challenges & Directions
Policies and Compliance in US Financial Industry Today

- Ubiquitous high-stakes Regulatory Compliance requirements
  - Sarbanes Oxley, SEC, HIPAA, etc.
- Internal company policies about access, confidentiality, transactions
  - For security, risk management, business processes, governance
- Complexities guiding who can do what on certain business data
- Often implemented using rule techniques
- Often misunderstood or poorly implemented leading to vulnerabilities
- Typically embedded redundantly in legacy silo applications, requiring high maintenance
- Policy/Rule engines lack interoperability
<table>
<thead>
<tr>
<th>Classification</th>
<th>Application</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merchant</td>
<td>Purchase Approval</td>
<td>If credit card has fraud reported on it, or is over limit, do not approve.</td>
</tr>
<tr>
<td>Mutual Funds</td>
<td>Rep trading</td>
<td><em>Blue Sky</em>: State restrictions for rep’s customers.</td>
</tr>
<tr>
<td>Mortgage Company</td>
<td>Credit Application</td>
<td>TRW upon receiving credit application must have a way of securely identifying the request.</td>
</tr>
<tr>
<td>Brokerage</td>
<td>Margin trading</td>
<td>Must compute current balances and margin rules before allowing trade.</td>
</tr>
<tr>
<td>Insurance</td>
<td>File Claims</td>
<td>Policy States and Policy type must match for claims to be processed.</td>
</tr>
<tr>
<td>Bank</td>
<td>Online Banking</td>
<td>User can look at own account.</td>
</tr>
<tr>
<td>All</td>
<td>House holding</td>
<td>For purposes of silo (e.g., statements or discounts), aggregate accounts of all family members.</td>
</tr>
</tbody>
</table>
Policies for Compliance and Trust Mgmt.: Role for Semantic Web Rules

• Trust Policies usually well represented as rules
  – Enforcement of policies via rule inferencing engine
  – E.g., Role-based Access Control
    • This is the most frequent kind of trust policy in practical deployment today.
  – W3C P3P privacy standard, Oasis XACML XML access control emerging standard, …

• Ditto for Many Business Policies beyond trust arena, too
  – “Gray” areas about whether a policy is about trust vs. not: compliance, regulation, risk management, contracts, governance, pricing, CRM, SCM, etc.
  – Often, authorization/trust policy is really a part of overall contract or business policy, at application-level. Unlike authentication.
  – Valuable to reuse policy infrastructure
Advantages of Standardized SW Rules

- Easier Integration: with rest of business policies and applications, business partners, mergers & acquisitions
- Familiarity, training
- Easier to understand and modify by humans
- Quality and Transparency of implementation in enforcement
  - Provable guarantees of behavior of implementation
- Reduced Vendor Lock-in
- Expressive power
  - Principled handling of conflict, negation, priorities
Advantages of SW Rules, cont’d:

Loci of Business Value

- Reduced system dev./maint./training costs
- Better/faster/cheaper policy admin.
- Interoperability, flexibility and re-use benefits
- Greater visibility into enterprise policy implementation => better compliance
- Centralized ownership and improved governance by Senior Management
- Rich, expressive trust management language allows better conflict handling in policy-driven decisions
Outline

• Introduction
  – Challenge of Semantics
  – Opportunities of the New Generation Web

• Policies and Compliance
  – Landscape Today
  – Advantages of Standardized Semantic Web Rules (Situated Courteous Logic Programs in RuleML)

• Financial Information Integration
  – Conflicting Definitions in Business Reporting
  – Mapping Approach (Extended COntext INterchange)

• Research Challenges & Directions
Equational Ontological Conflicts in Financial Reporting

# of customers = # of end_customers + # of distributors

Gross Profit = Net Sales – Cost of Goods

P/E Ratio = Price / Earnings(last 4 Qtr)

Price = Nominal Price + Shipping

# of customers = # of end_customers + # of prospective customers

Gross Profit = Net Sales – Cost of Goods – Depreciation

P/E Ratio = Price/ [Earnings(last 3 Qtr) + Earnings(next quarter)]

Price = Nominal Price + Shipping + Tax

“heterogeneity in the way data items are calculated from other data items in terms of definitional equations”
Solution Approach: ECOIN
Extended COntext INterchange   MIT Sloan prototype
E-Shopping App. (Financial Info is ubiquitous in e-biz)

Query
Prices of Products
Cheaper in eToys

eToys

Kid’s World

Context Mediator

Price Equations

Pokemon

Price: Nominal + Tax + Shipping
Product Code: Alpha

Pokemon

Price: Nominal
Product Code: Numeric

Pokemon

17

Pokemon

20

Pokemon

234567

5/18/2004 Copyright 2004 by Benjamin Grosos. All Rights Reserved
Outline

• Introduction
  – Challenge of Semantics
  – Opportunities of the New Generation Web

• Policies and Compliance
  – Landscape Today
  – Advantages of Standardized Semantic Web Rules (Situated Courteous Logic Programs in RuleML)

• Financial Information Integration
  – Conflicting Definitions in Business Reporting
  – Mapping Approach (Extended COntext INterchange)

• Research Challenges & Directions
Research Challenges: Core

• Integrating rules with ontologies
  – Rules refer to ontologies (e.g., in RuleML)
  – Rules to specify ontologies (e.g., Description Logic Programs)
  – Rules to map between ontologies (e.g., ECOIN)
  – Combined rules + ontologies knowledge bases (e.g., RuleML + OWL)

• Describing business processes & web services via rules + ontologies
  – Rules query web services (e.g., in RuleML Situated feature)
  – Rules trigger actions that are web services (e.g., ditto)
  – Capture object-oriented process ontologies
    • Default inheritance via rules (e.g., Courteous Inheritance)
    • Wrapper/transform to legacy C++, Java, UML
    • Develop open source knowledge bases (e.g., MIT Open Process Handbook Initiative)
  – Event triggering of rules (e.g., capture ECA rules in RuleML)
Research Challenges: Business Policies

- Apply advanced rule and ontology representation to business policies in compliance, trust, contracts, etc.
  - Application scenarios for compliance checking/support services intra- and inter-enterprise
  - Policy language & engines on top of rule language & engines
  - In/with existing/emerging standards: XBRL, XACML, P3P, ebXML, EDI, Legal XML, …
  - Strategy and roles in the market ecology: regulators, communal repositories, service providers, etc.
  - Embedding into the bigger pictures of financial services, e-commerce, semantic web services, business process automation
Context: Our Overall Research Agenda

- Invent Core Technologies and concepts of the New Generation Web
  - Semantic Web; Rules and RuleML emerging standard
    - supporting knowledge representation theory of Situated Courteous Description Logic Programs
  - Semantic Web Services; Business Process Automation for B2B and EAI
    - Requirements analysis
    - Use of Rules, together with ontologies – or to represent ontologies

- Pilot Business Application Scenarios
  - End-to-end e-contracting, e.g., in manufacturing supply chain
    - SweetDeal approach using rules (plus ontologies)
  - Financial information and reporting:
    - ECOIN approach mapping ontologies
  - Other: security authorization, travel, …

- Analyze Prospective Early Adopter Areas
  - Strategy: Adoption Roadmap; Market Evolution
  - Entrepreneurial Opportunities
More Strategic Opportunities in Compliance

• XBRL (eXtensible Business Reporting Language):
  – SWS rules + ontologies can reduce degree of industry consensus required to enable interoperability
    • Difficult to get agreement on single definition of “earnings”; easier to agree on “long-term capital gains realized from sale of real estate assets”.
    • Translate between different use contexts’ ontologies

• SEC and other regulatory agencies:
  – They can accelerate compliance
    • via providing automated SWS specifications of regulations and reporting forms (+ the instructions)
      – e.g., RuleML regulatory rulebases accessible via Web Services interfaces
Outline

• Introduction
  – Challenge of Semantics
  – Opportunities of the New Generation Web

• Policies and Compliance
  – Landscape Today
  – Advantages of Standardized Semantic Web Rules
    (Situated Courteous Logic Programs in RuleML)

• Financial Information Integration
  – Conflicting Definitions in Business Reporting
  – Mapping Approach (Extended COntext INterchange)

• Research Challenges & Directions
Quickie Bio of Presenter

- MIT Sloan professor since 2000
- 12 years at IBM T.J. Watson Research; 2 years at startups
- PhD Comp Sci, Stanford; BA Applied Math Econ/Mgmt, Harvard
- Semantic web services is main research area:
  - Rules as core technology
  - Business Applications, Implications, Strategy:
    - e-contracting/supply-chain; finance; trust; …
  - Overall knowledge representation, e-commerce, intelligent agents
- Co-Founder, Rule Markup Language Initiative — the leading emerging standards body in semantic web rules (http://www.ruleml.org)
- Core participant in Semantic Web Services Initiative — which coordinates world-wide SWS research and early standards (http://www.swsi.org)
  - Area Editor for Contracts & Negotiation, Language Committee
  - Co-Chair, Industrial Partners program (SWSIP)
Semantic Web: concept, approach, pieces

- Shared semantics when interchange data \implies \text{knowledge}
- Knowledge Representation (cf. AI, DB) as approach to semantics
  - Standardize KR syntax, with KR theory/techniques as backing
- Web-exposed Databases: SQL; XQuery (XML-data DB’s)
  - Challenge: share DB schemas via meta-data
- RDF: “Resource Description Framework” W3C proposed standard
  - Meta-data lower-level mechanics: unordered directed graphs (vs. ordered trees)
  - RDF-Schema extension: simple class/property hierarchy, domains/ranges
- Ontology = formally defined vocabulary & class hierarchy
  - OWL: “Ontologies Working Language” W3C proposed standard
    - Subsumes RDF-Schema and Entity-Relationship models
    - Based on Description Logic (DL) KR ~subset of First-Order Logic (FOL))
- Rules = if-then logical implications, facts ~subsumes SQL DB’s
  - RuleML: “Rule Markup Language” emerging standard
    - Based on Logic Programs (LP) KR ~extension of Horn FOL
Some Semantic Web Advantages for Biz

• Builds upon XML’s much greater capabilities (vs. HTML*) for **structured detailed descriptions** that can be processed **automatically**.
  – Eases application development effort for **assimilation of data in inter-enterprise interchange**

• Knowledge-Based E-Markets -- where Agents Communicate
  (Agent = knowledge-based application)
  – ⊃ potential to **revolutionize interactivity in Web marketplaces**: B2B, …

• Reuse same **knowledge for multiple purposes/tasks/app’s**
  – Exploit declarative KR; Schemas

• * new version of HTML itself is now just a special case of XML
Web Service -- definition

(For purposes of this talk:)

- A procedure/method that is invoked through a Web protocol interface, typically with XML inputs and outputs
  - Add the flexibility of XML to the concepts of RPC
  - XML Tools support extra functionality required

- Purpose: Program integration across application and organizational boundaries
  - Needs commercial semantics
Semantic Web Services

- Convergence of Semantic Web and Web Services
- Consensus definition and conceptualization still forming
- Semantic (Web Services):
  - Knowledge-based service descriptions, deals
    - Discovery/search, invocation, negotiation, selection, composition, execution, monitoring, verification
    - Advantage: **reuse** of knowledge across app’s, these tasks
      - Integrated knowledge
  - (Semantic Web) Services: e.g., infrastructural
    - Knowledge/info/DB integration
    - Infererencing and translation
B2B Tasks: Communication for Business Processes with Partners

• B2B business processes involving significant Communication with customers/suppliers/other-partners is overall a natural locus for future first impact of SWS.

• Customer Relationship Management (CRM)
  – sales leads and status
  – customer service info and support

• Supply Chain Management (SCM):
  – source selection
  – inventories and forecasts
  – problem resolution
  – transportation and shipping, distribution and logistics

• orders, payments, bill presentation
Some B2B Tasks (continued)

- bids, quotes, pricing, **CONTRACTING; AUCTIONS;** procurement
- authorization (vs. authentication) for credit or trust
- database-y: e.g.,
  - catalogs & their merging
  - policies
- inquiries and answers; live feedback
- notifications
- trails of biz processes and interactions
- ratings, 3rd party reviews, recommendations
- knowledge management with partners/mkt/society
Some Answers to:
“Why does SWS Matter to Business?”


• 2. “Business processes require communication between organizations / applications.” - Data and programs cross org./app. boundaries, both intra- and inter- enterprise.

• 3. “It’s the automated knowledge economy, stupid!” - The world is moving towards a knowledge economy. And it’s moving towards deeper and broader automation of business processes. The first step is automating the use of structured knowledge.
  – Theme: reuse of knowledge across multiple tasks/app’s/org’s
Role of Standards

• Obs.: **Standards** are crucial, and central, to integration in an open era.

• → high percentage of effort invested in standards development in new generation web (XML, WS, SW, SWS)

• In SWS, this begins with basic research!

• Lots of strategy surrounding standards.

• Emerging standards efforts include much research.