Information Flow in the Federal Enterprise Redux

richard.murphy@gsa.gov

What's the General Services Administration ?

GSA leverages the U. S. government's buying power to help Federal agencies better serve the public with:

* space (Federal court houses, port of entry and border stations, future work space, green buildings)

* stuff (supply chain, logistics, disaster response, green technologies)

* solutions (government wide acquisition contracts, federal acquisition regulations, office of government-wide policy, office of citizen services and communications)

Four key public policy issues:

1. Create efficient marketplace for emerging technologies by complementing standards with knowledge like Federal Enterprise Architecure Reference Model Ontology (FEA-RMO) and US Privacy Act of 1974 OWL-DL ontologies

2. Optimize risk transfer in Federal acquisition through performance based contracting with required trade-offs among databases, rules engines, reasoners and theorem provers

3. Citizen Privacy Service, an OSERA reference implementation work-in-progress, J2EE smart-bus middleware with policy decision point (privacy entailment with Pellet) and policy enforcement point (FOL provenance in JTP)

4. Information sharing based on Information Flow: The Logic of Distributed Systems, Goguen's Theory of Institutions and automated ontology mapping with Hasse language maps and Chu Spaces

Information Flow as Metaphor

"The espionage ring tone carries the information that Ginny was the person calling Rick." More generally, we claim that a's being of type α carries the information that b is of type β

Information Flow Principles

First Principle (P1): *IF* results from regularities in a distributed system

Second Principle (P2): *IF* crucially involves both types and their particulars (tokens)

Third Principle (P3): It is by virtue of regularities among connections that information about some components of a distributed system carries information about other components

Fourth Principle (P4): The regularities of a given distributed system are relative to its analysis in terms of information channels

<snip, hope you read the position paper ...>

Information Flow Scenario - Languages, Logics, Models and Theories

Three classifications representing local and remote components in the distributed system are: 1) a set of Unified Modeling Language (UML) artifacts held in a Meta Object Facility (MOF) at DIA, 2) a set of Web Ontology Language (OWL) artifacts held in a triple store at DHS, and 3) a set of XML Topic Maps (XTM) held in a meta-data repository at DOJ.

Figure 1 - Three Classifications Representing Local and Remote Components

MOF	Triple Store	Repository
Σ_M	Σ_T	Σ_R
$\vert\models_M$	$\vert\models_T$	$\vert\models_R$
M	T	R

Description Logic Expressiveness[1]

1. "A Description Logic for Use as the ODM Core," Lewis Hart and Patrick Emery, EDOC 2004

Nomen- clature	Meaning	DL-Core SHIN(D)	UML ALHOIN(D)	OWL-DL SHOIN(D)	ER ALN(D)	Topic Maps AL(D)
AL	Atomic Concept Universal Concept Bottom Concept Atomic Negation Intersection Value Restriction Limited Existential Quantification	x	Atomic Concept Value Restriction Limited Existential Quantification (AL-)	x	Atomic Concept Value Restriction Limited Existential Quantification (AL-)	Atomic Concept Value Restriction (AL - -)
C	Full Negation or Complement	X		X		
E	Full existential Quantification	x		x		
H	Role Hierarchies	x	x	x		
I	Inverse Roles	x	x	x		
N	Unqualified Number Restrictions	x	x	x	x	
O	Enumerated Classes			x		
R*	Transitive Roles	x		x		
V	Union Constructor	x		x		
(D)	Datatypes	x	x	x	x	x

<snip ...>

Institutions - preserving satisfiability under non-monotonic reasoning

Definition 1. *An Institution consists of an abstract category* Sign*, the objects of which are signatures, a functor* Sen: $\mathsf{Sign} \to \mathsf{Set}$*, and a functor* Mod: $\mathrm{Sign}^{\mathrm{op}} \to$ Set*. Satisfaction is then a parameterized relation* $\models_\Sigma$ *between Mod(*Σ*) and Sen (*Σ*), such that the following condition holds for any signature morphism* $\psi\colon \Sigma \to \Sigma'$*, any* Σ'*-model* M'*, and any sentence e,* $M' \Vdash_{\Sigma'} \psi(e)$ iff $\psi(M') \Vdash_\Sigma e$ *where* $\psi(e)$*abbreviates* Sen*(*ψ*)*(e) *and* $\psi(M')$ *abbreviates* Mod$(\psi)(e)$.

Where infomorphisms allow for structure preserving transformation, institution morphisms preserve satisfiability through semantic preserving transformation. Structure preserving transformation can result in a loss of decideability. Consider the transformation from DIA's UML MOF to DHS's OWL triple store. A UML (Class, Association, Class) structure classifies as ALHOIN(D) in description logi. This structure transforms to the OWL structure (Class,Property,Class), classified in description logic as OWL Full.

Figure 2. Semantic preserving transformation between DIA and DHS.

$$\begin{array}{ccc} M\ \text{Sentences} & \xrightarrow{I} & \text{T Sentences} \\ \Big| \models_M & & | \models_T \\ \text{M Structures} & \xleftarrow[I]{} & \text{T Structures} \end{array}$$

<snip ...>

Information Channels and Living Strucure (aka. Living Liasing Languages)

Definition 2. *An information channel consists of an indexed family $C = \{ f_i : A_i \rightleftarrows C \}_{i \epsilon I}$ of institution morphisms with a common codomain C, called the core of the channel.*

In addition to allowing DOJ, or any member of the federation, to participate in one channel, the Natural Laws of Federation, or complex systems theory, suggest that we need may need more than one channel, or that we allow members to change channels, to more effectively share information.

Semantic Information Services - One more point of differentiation ...

Description Logic (Pellet) provides classification, realization, consistency, satisfiability

First Order Logic (SNARK) provides skolemization, unification, resolution, generalized modus ponens

Information Sharing Futures

Algebraic specifications, Institutions, semiotics, connections, unskolemization, category theory, formal concept analysis, information flow framework, hasse language maps, Chu spaces