19th ICCRTS

Title of Paper: Semantic Information Management Control of Mission Asset State Changes

Assigned Paper #: 026

Primary Topic: Organizational Concepts and Approaches

Alternatives: Modeling and Simulation;
Data, Information, and Knowledge

Authors

Jason Bryant Gregory Hasseler Matthew Paulini Tim Lebo

Air Force Research Laboratory
RISA - Operational Info Management
525 Brooks Road
Rome, NY 13440

Point of Contact: Jason.Bryant.8@us.af.mil 315-330-7670

Organization:
Air Force Research Laboratory - RISA
Operational Information Management Branch
525 Brooks Road
Rome, NY 13441

Organization Phone: 315-330-7670

Organization E-mail: Jason.Bryant.8@us.af.mil

Distribution A: Approved for public release; distribution is unlimited.

(Approval AFRL PA # 88ABW-2014-1864)

Semantic Information Management Control of Mission Asset State Changes

Jason Bryant Gregory Hasseler Matthew Paulini Tim Lebo
Air Force Research Laboratory - RISA Information Management Technologies
Distribution A: Approved for public release; distribution is unlimited.
(Approval AFRL PA #88ABW-2014-1864)

ABSTRACT
Managing state changes for semantically represented mission assets can be complex and burdensome. Standard practice simplifies updating state-based relationships by performing sequential delete and add operations, or by creating a new instance of the asset with updated state values. The former is inflexible and constraining as it eliminates historical state tracing, making it impossible to implement state provenance and advanced semantic queries. The latter causes mass duplication of non-stateful attributes while requiring asset redefinitions upon each stateful relationship change event. Our research implements an alternative solution by representing the asset definition once and declaring each state change as an instance specialization under the W3C Provenance Ontology (Prov-O). This minimizes duplication and cleanly disentangles stateful attribute changes from the administration of semantic asset instantiation. This approach conforms to practical mission asset models, such as Blue Force Tracking, treating state changes as iterative, independent updates. The minimal overhead consists of setting a valid start time for the state specialization. These enhancements to modeling stateful mission relationships enable a more agile and discoverable awareness to operational clients. It increases agility within C2 Information domains by supporting resource introspection for access control, message prioritization, and produces capabilities for historical state-based asset queries.

1.0 INTRODUCTION

Enterprise and degraded C2 Information domains have use cases for semantics when the feature set is rich enough to meet application and modeling requirements. Managing relationships that can change over time can be inelegant, limited, or overly burdensome for semantic technologies. Historically, the use cases that semantics have been applied to are oriented toward document management (e.g. web) or the representation of a particular domain of knowledge, such as time, social associations, or geospatial coordinates. For semantic relationships that involve churning values, whether that turnover is regular or irregular, there are no semantic standard or technology solutions that are sufficient.

Certain facets of the standards often make it challenging to solve problems in a real world, operational setting. Creating solutions for challenges within the semantic knowledge representation, link analysis, query expressions, query results, and natural language processing domains has been the focus of existing research, resulting in reification, standards-based provenance, and ontology versioning. Reification can be solved through enhancing ontologies to support finer grained relationships and concepts, so that the point is never reached where reification becomes necessary. Provenance can, in a basic form, be overcome through the adoption of named graphs instead of pure triples, thus enabling a simple but foundational level of document sourcing and traceability. Resolving conflicting versions of ontology entities and predicates is overcome through adoption of standard Source/Domain/Version (e.g. http://mysourcenamespace.com/stateontology/V2.0/) structures for URIs. While these approaches are relatively simple to implement and effective in providing solutions to those challenges, they can increase complexity, decrease maintainability and performance, and do not provide a foundation for furnishing more sophisticated features in semantic management.

A challenge that has remained resistant to previous efforts is management of entity data, object states, and their transitions to new values. In practice, semantic state representation challenges are a side effect of integrating semantic technologies with distinct areas of focus; one being knowledge representation (RDF), the other being ontology instantiation (OWL). OWL can be seen as the semantic web equivalent of schemas to the standardized document object model. The real-world semantic "primitives" are ontological entities, such as a person, place, concept, or mission asset. These ontological entities are essentially the serialization of a triple-set. They possess a fixed identity throughout their lifecycle, comparable to semantic URIs, but link to data attributes that are not necessarily fixed. These stateful data attributes need to be represented either ontologically as an entity instance or as an unstructured relational fact defined without an OWL counterpart.

Applying these observations to an operational, real world use case demonstrates the complexities of mapping abstract solutions to concrete semantic problems. Asset examples include entities such as a document, an F15, an Air Tasking Order (ATO) knowledge extraction, or a Blue Force Tracking (BFT) message. Semantic facts expressed within these entities could be a point of contact name, current geo-location status, fuel status, social POCs, or target details. These examples result in two alternate operational semantic state management approaches, which are described below.

One approach is that only a single semantic instance of an asset exists, and it is updated with the 'present' relationship values whenever they are determined. Implementing this operational view is simplistic and results in a constantly up-to-date semantic model. However, it excludes all past state changes for an asset's dynamic relationships, such as geolocation, role, operational condition, or health. This view is seriously flawed for most operational domain use cases. By replacing old state values with new static values, it removes any capacity for tracking non-static relationship values. The view is always restricted to the 'present', meaning history can never be queried and, consequently, never learned from.

The second possible approach maintains a historical record by creating a new instance, or concretization, of the asset and its present relationship states. Over time, particularly if an asset has relationships that change quite often, thousands of versioned instances of an asset could be created, making queries overly complex and resulting in a high degree of duplication for relationships that stayed the same, yet are not static in nature. As an example, consider a semantic representation of an F15 with its various possible relationships, such as tail number, vehicle type, fuel supply, speed, direction, geospatial location, pilot, call name, and radio frequency. Some of these relationships are static, or at least change infrequently, and yet would need to be re-declared as relationships for each triggered new instance of the F15. One of our use cases for Blue Force Tracking calls for a sensor-generated state change message to be sent every 30-45 seconds. A single day of semantic extractions could result in 2,880 new instances of the asset, each with many duplicated relationships. While this view maintains traceability, enabling semantic queries of both present and past relationships, it also overburdens storage and inferencing resources, and overcomplicates the query process by treating state changes as duplication triggers. All queries for an asset would be required to disentangle past and current instances.

The specialization of semantic assets we propose avoids the pitfalls of previous approaches by utilization of an ontology with the appropriate vocabulary to manage stateful relationships of semantic assets. This approach leads to simplified modeling, is more harmonious with object instantiation rules, and creates a foundation over which a set of advanced capabilities can be developed.

2.0 MOTIVATION AND RELATED WORK

As semantic standards mature and applications expand into new domains, research regarding semantic management of stateful relationships is beginning to be explored more fully. Current research has been tangential at best, while missing many of the niche problem areas of semantics. Approaches in this area have focused on inferencing by using join sequences (Yang, 2013) or resolving models with conflicting states (Zhang, 2009). To wit, approaches involving applied analytics for state management have attempted to do so during the extraction phase of data processing (Turney and Pantel, 2010), rather than utilizing semantic technologies or ontology models.

Research in state management for semantics requires additional, targeted explorations with supplementary resources, particularly as semantic applications and their domains grow to include more complex ontologies and object models.

Our approach seeks to solve a combination of challenges within Information Management (IM), Semantic Information Modeling, Data to Intelligence (D2I), and Information Retrieval (IR).

3.0 STATE MANAGEMENT OF INFORMATION AND MISSION ASSETS

Semantic technologies have generally been used to represent the categorization and properties of a domain of knowledge, data extracted from documents, or to represent information over which some form of rule logic and learning can be executed. For most of these cases, there exists no notion of state. Additionally, most semantically represented bodies of knowledge involve the declaration of data properties that tend to be evaluated on their existence rather than on their propensity to change. There are general assumptions of statelessness or non-stochastic truth. For the rule logic use case, the focus is on inferencing and reasoning for truth determinations or, at a minimum, to monitor events. In the most standard use case, that of content-based document modeling (World Wide Web paradigm), documents change infrequently and, when they do, can be considered a new version of the old document. For some documents, this makes sense (e.g. Reports, iteratively improved content, or books). For others, the only association between old content and new content may be the source URI (e.g. Newspaper, Facebook news wall, search engine front end). In operational use cases where mission assets, people, and targets are involved, semantic asset representation is a different beast with a set of unique challenges.

Semantic assets, similar to documents, have an intrinsic identity with possessive traits. The distinct difference, however, is that the assets are much more likely to have attributes that change state, which should never reflectively alter the identity of the possessive asset. For example, consider a newspaper as an iterative publication with the same title at the top each day, but which is fundamentally a new instance with static contents. In the case of a web-based newspaper, there is no temporally canonized version, rather a fluctuating set of aggregated news stories throughout the day. Also, the attributes a mission asset possesses are much more likely to be inter-related to other asset instances, whereas the focus of many document-centric extractions are tagging, keyword frequency analysis, or knowledge representation.

Traditional use cases emphasize an iterative, document-centric versioning paradigm, while our approach emphasizes overlays with independently managed assets containing stateful data and object properties. The key relationship used for this is the `specializationOf` predicate of the Provenance Ontology (Prov-O) W3C recommendation. It is intended to apply state-based relationships to any Entity, Agent, or Activity, auspices under which any semantic asset instance should fall. The example below illustrates the difference between a semantically defined mission asset and its target status specialization is drawn from an operational scenario with Blue Force Tracking, ATO, Red Force Tracking, Close Air Support (CAS) Requests, Intelligence Reports, and Battle Damage Assessments.

4.0 EXPERIMENT: USE CASE

Creating experiment support for testing applied use cases for our semantic state management approach required building an Information Management (IM) infrastructure. This included establishing sample publishers and consumers, a quad-store (Parliament), a raw document database (Hash Map with file references), and other integrated IM technologies. The distinct phases of data processing are illustrated in Figure 1.

Simulated mission data was published and semantically expressed via a set of indexers. The result of semantic processing is a semantic RDF/OWL document that relates values for details involving times, locations, missions, targets, points of contact, etc. Figure 2 shows an RDF/OWL document resulting from extracting mission details from an Air Tasking Order. The full ATO document is expansive, with additional content for targets, mission codes, and more.

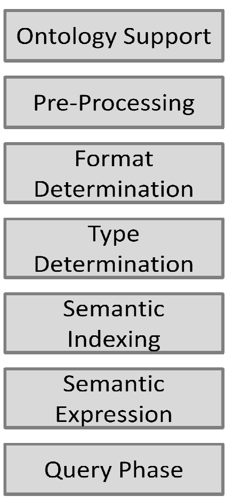

Figure 1 - Process Flow

Each document is viewed concurrently as an independent publication with associated semantic named graph. Each named graph has a minimum of the following fields:

- Named Graph URI
- Information URI
- Publisher Identity URI
- Publisher Role
- Message Topic
- Message Type
- Message Format
- Time Published

Optional fields exist for a wide range of semantic relationships, including:

- Mission Involvement
- POC Involvement
- Asset Involvement
- Target Involvement
- GeoSPARQL Compatible Geolocations
- Keywords
- Publisher Geolocation

Following the establishment of the supporting infrastructure, appropriate semantic models for object instances needed to be adopted. Ontologies supported by our framework include the common solutions for time, geospatial (GeoSPARQL), common elements (U-Core SL), and mission planning (Cornerstone), with a custom ontology for information management. Additionally, before the semantics can be extracted there is a pre-processing stage consisting of format determination and xml type determination(e.g. schema, CoT message type, MTF-XML type), if applicable. The Aperture open source project was adopted to provide the majority of this solution, although some customization for DoD formats was required. After the pre-processing (i.e. format determination, type determination, and semantic extraction) completes, state-based

```
<http://anchor/cornerstoneCore#Mission?id3535>
  rdf:type cornerstoneCore:Mission ;
  cornerstoneAir:hasSortie
       [ rdf:type cornerstoneAir:AirSortie ;
        cornerstoneAir:sortieAircraft
            [ rdf:type cornerstoneAir:F15E ;
             cornerstoneCore:name
                 "F15E" ;
             extension:quantity
                 [ rdf:type cornerstoneCore:Quantity ;
                  cornerstoneCore:hasNumericValue
                      "2"^^<http://www.w3.org/2001/XMLSchema#float>
                 ]
            ]
       ] ;
  cornerstoneCore:missionKey
       "3535" ;
  cornerstoneCore:pocInfo
       [ rdf:type cornerstoneCore:POCInfo ;
        cornerstoneCore:poc
            [ rdf:type cornerstoneCore:Agent ;
             cornerstoneCore:jobTitle
                 "LT COL" ;
             cornerstoneCore:memberOf
                 <http://anchor/cornerstoneCore#MilitaryUnit?id522_FS> ;
             cornerstoneCore:name
                 "FRANCIS"
            ]
       ] ;
  cornerstoneCore:primaryMissionCode
       "CMD" ;
```

Figure 2 - Semantic Mission Snippet

queries are enabled for any new observations of state. Sample semantic snippets of a target is in Figures 3 and the basic metadata of a mission report in Figure 4.

```
cornerstoneCore:requestedTarget
      [ rdf:type im:Target ;
       cornerstoneCore:description
           "ANTMAST" ;
       cornerstoneCore:id "0992-03815DD001" ;
       cornerstoneCore:location
           [ rdf:type cornerstoneCore:GeographicPoint ;
            geo:lat "36.71472"^^<http://www.w3.org/2001/XMLSchema#float> ;
            geo:long "-119.548615"^^<http://www.w3.org/2001/XMLSchema#float>
           ] ;
       cornerstoneCore:name
           "SANGER3RDZONEOPSCENTER"
      ] ;
```

Figure 3 - Target Semantic Snippet

The mission-based test scenario consisted of 230 messages published over a period of 10 minutes. The message types included USMTF formatted ATOs, Intel Reports, Battle Damage Assessments, and Close Air Support Requests, as well as Cursor-on-Target formatted Blue Force Tracking messages for different mission assets such as F15s, UAVs, MRAPs, and JTAC Red Force. The semantic relationships created were produced by means of the extraction framework we created in our previous ICCRTS research (Bryant and Paulini,

```
<http://phoenix.rl.af.mil/im.owl#Information?id00c135fb0fac5e124c619f09660a>
   rdf:type im:Information ;
   extension:aboutMission
        <http://anchor/cornerstoneCore#Mission?id2105> ;
   extension:hasMissionReport
        <http://phoenix.rl.af.mil/extension.owl#MissionReport?id2105> ;
   im:hasPayload
<http://phoenix.rl.af.mil/im.owl#Payload?id18232c929fe7405f2c01abd1fb27> ;
   im:hasPublisher <http://phoenix.rl.af.mil/scenario/1.0/ASOC_OODO> ;
   im:informationType "mil.af.rl.xml-mtf" .

<http://phoenix.rl.af.mil/scenario/1.0/ASOC_OODO>
   im:publishedInformation
        <http://phoenix.rl.af.mil/im.owl#Information?id00c135fb0fac5e124c619f09660a> ;
   im:publisherId "ASOC_OODO" .

<http://phoenix.rl.af.mil/im.owl#Payload?id18232c929fe7405f2c01abd1fb27>
   rdf:type im:Payload ;
   im:formatId "xml-mtf" ;
   im:type "MISREP" .

<http://anchor/cornerstoneCore#Mission?id2105>
   rdf:type cornerstoneCore:Mission .

<http://phoenix.rl.af.mil/extension.owl#MissionReport?id2105>
   rdf:type extension:MissionReport ;
   extension:aboutMission
        <http://anchor/cornerstoneCore#Mission?id2105> .
```

Figure 5 - Intel Report Metadata

2013), although there is now added support for GeoSPARQL-based location extractions.

After publication, the quad-store has instances defined for all missions and assets involved, but no continuant state changes. Figure 5 expresses the semantic expression of a state within a published Battle Damage Assessment report, showing how to utilize our state management approach.

```
:target_antmast_0992-03815DD001
     prov:specializationOf   :target_antmast ;
     prov:startedAtTime "2012-04-12T00:00:00-04:00"^^xsd:dateTime ;
     cornerstoneCore:damageState "damaged" .
```

Figure 6 - Battle Damage Assessment Status Change

An advanced set of query capabilities can be enacted over this data. This allows state traceability and discoverability of the newest particular state of an asset. These are key features that were previously not available semantically. Figure 6 illustrates a SPARQL query for the last 50 condition states of a target, while Figure 7 makes a few slight modifications to find only the newest specialization involving the target's damaged state.

```
PREFIX prov: <http://www.w3.org/ns/prov#>

SELECT *
WHERE {
      ?s prov:specializationOf prov:target_antmast .
      ?s prov:startedAtTime ?time . }
ORDER BY DESC(?time)
LIMIT 50
```

Figure 4 - Last 50 Status Changes for Target

```
PREFIX prov: <http://www.w3.org/ns/prov#>

SELECT *
WHERE {
      ?s prov:specializationOf prov:target_antmast .
      ?s cornerstoneCore:damageState ?o .
      ?s prov:startedAtTime ?time . }
ORDER BY DESC(?time)
LIMIT 1
```

Figure 7 - Newest Damage Assessment

5.0 CONCLUSION

Traditional semantic data model approaches fall short when confronting the challenge of state-based relationships. They focus on static knowledge representation, extractions of static data properties, or enabling of information management features via rule engines and inferencing. Our use case overlays the traditional approaches with a layer of independently defined mission assets to reconcile integration conflicts. Managing states for data and object properties are not unique to mission-based or operational-centric assets, but applicable to all stateful semantic resources.

We theorize that managing the state of assets is significant in reducing the computation time of semantic queries, the load on semantic DBs, and eliminating wasteful property and instance duplications. Managing states appropriately also enables advanced query heuristics, makes state change instances more lightweight, and organizes state changes temporally so that the newest events are easier to discover.

In our future work, we explore semantic state traceability paired with semantic graph analytics. Reasoning over stateful trends within segmented time periods can demonstrate possible advanced uses of semantics for stochastic and Boolean-based analytics applications, thus producing support for prioritization, query result set ordering, and provenance modeling of analytics.

REFERENCES

Turney, P., Pantel, P., "From Frequency to Meaning:Vector Space Models of Semantics", Journal of Artificial Intelligence Research, 2010.

NATO Research Task Group SAS-085, "Executive Overview of C2 Agility". STO-TR-SAS-085, STO Technical Report, http://www.dodccrp.org/sas-085/sas-085_report_overview.pdf.

Bunescu, R., Mooney, R., "Collective Information Extraction with Relational Markov Networks". 42nd Annual Meeting of the Association for Computational Linguistics, July, 2004.

Bryant, J., Combs, V., Hanna, J., Hasseler, G., Krokowski, T., Lipa, B., Reilly, J., Tucker, S., "Phoenix Base Implementation: Final Technical Report", DTIC, 2010.

Barroso, L.A., Dean, J., Holze, U., "Web Search for a Planet: The Google Cluster Architecture", Micro, IEEE: Volume 23: Issue 2 (pp. 22 - 28), 2003.

Zhang, X., "An Ensemble Learning and Problem Solving Architecture for Airspace Management", Applications of Artifiial Intelligence Conference, 2009.

Yang, C., "A Formal Semantics for Stateful Aspect Weaving", Journal of Information and Computer Sciences, 10:8, 2013.

Bryant, J., Combs, V., Hanna, J., Hasseler, G., Hillman, R., Lipa, B., Reilly, J., Vincelette, C., "Phoenix: An Abstract Architecture for Information Management Final Technical Report", DTIC, 2010.

Bryant, J., Paulini, M., "Making Semantic Information Work Effectively for Degraded Environments", ICCRTS, 2013.

Schrage, M., "The Struggle to Define Agility", CIO Magazine, August 2004.

Shulstad, R.A., "Cursor on Target. Inspiring Innovation to Revolutionize Air Force Command and Control", Air & Space Power Journal, winter 2011, http://www.airpower.au.af.mil/apjinternational/apj-c/2012/2012-2/2012_2_08_shulstad-E.pdf.

Sudo, K., Sekine, S., Grishman, R., "An Improved Extraction Patterns Representation Model for Automatic IE Pattern Acquisition". 41st Annual Meeting of the Association for Computational Linguistics, July, 2003.

"Command and Control Joint Integrating Concept", Final Version 1.0, DTIC, http://www.dtic.mil/futurejointwarfare/concepts/c2_jic.pdf, 1 Sep 2005.

Lebo, T., Graves, A., McGuinness, D., "Content-Preserving Graphics", Consuming Linked Data Conference, 2013.

Tous, R., Delgado, J., "A Vector Space Model for Semantic Similarity Calculation and OWL", IN DEXA, 2006.