Towards a Formal Pedigree Ontology for Level-One Sensor Fusion

Christopher J. Matheus

Versatile Information Systems, Inc.

David Tribble

Referentia Systems, Inc.

Mieczyslaw M. Kokar

Northeaster University

Marion Ceruti and Scott McGirr

Space and Naval Warfare Systems Center

Outline

- Objectives
- The Web Ontology Language OWL
- Pedigree as Metadata
- Proposed Pedigree Ontology
- A Candidate Application
- Challenges

Our Objective

- Preliminary work to exploit data pedigree information to improve the interpretation and use of Navy track data (OTH GOLD)
  - develop an OWL ontology so we can formally reason about pedigree information
- Part of a larger ONR SBIR effort
  - to lessen information overload and improve the exchange of disparate data
  - uses a C2IEDM-based Track Ontology (OWL) for data mediation (see 11:00am talk today in Old Dominion)
  - joint work by VIS, Referentia Systems, Northeastern University and SPAWAR

Web Ontology Language

OWL is

- Semantic Web Ontology Language
- a W3C Recommendation (i.e., a standard)
- a formal language for writing ontologies;
  an ontology defines the language of discourse for a particular domain by identifying
  - pertinent object classes
  - relations between objects (e.g., object-properties)
  - object attributes (e.g., data-properties)
- developed by the W3C out of the DARPA-funded DAML project

OWL Pros and Cons

- Pros:
  - W3C standard with an XML syntax
  - well grounded formal semantics backed by years of research in formal logic, knowledge representation & artificial intelligence
  - growing number of applications for editing, checking and processing OWL documents
- Cons:
  - everything is represented with RDF triples
    - very low level and cumbersome to read and write
  - OWL properties are limited to binary relations
  - unable to represent general implications e.g. uncleOf(X,Y)

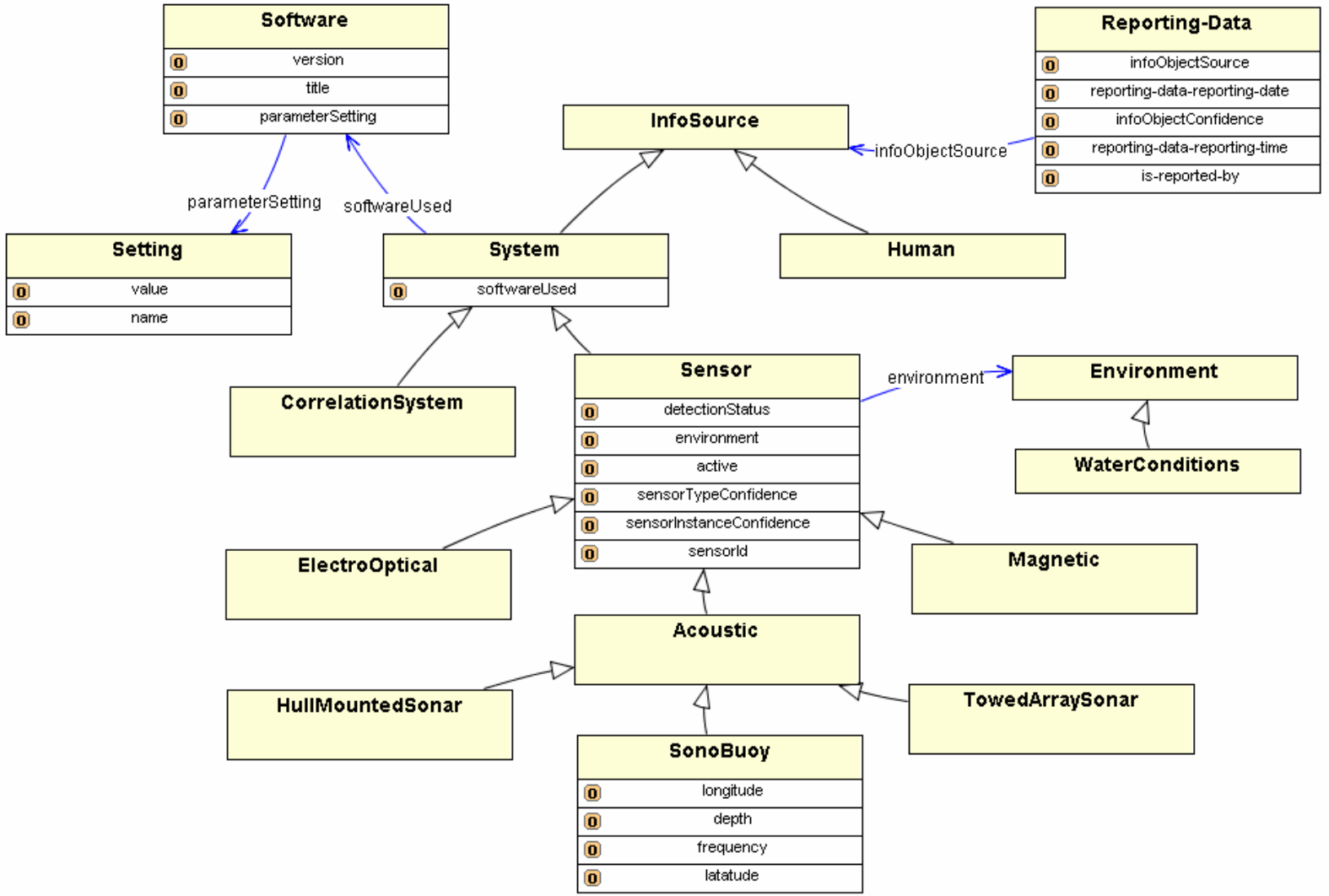
Pedigree = Metadata

- Data Pedigree = “data provenance” = “data lineage” = metadata = data about data
- Can include information about data such as
  - where did it come from?
  - what were the conditions it was derived under?
  - when was it sensed? received? expires?
  - how was it derived?
  - who produced it? owns it? validates it?
- Usually not available beyond simple time and source information
- Yet, can be important for evaluating quality, confidence, trust, usability, share-ability, etc.

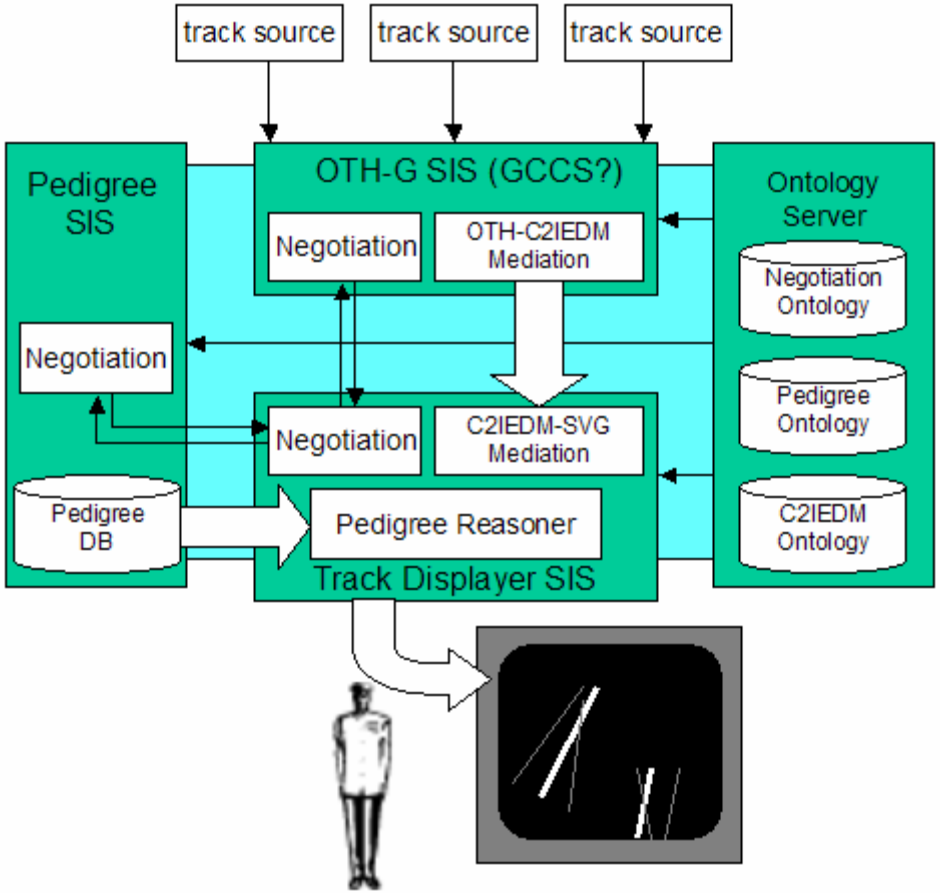
Track Data Pedigree

- OTH GOLD track data can provide information about provider, time & sensor type
- Other track data pedigree candidates:
  - sensor configuration
  - sensor type trust, sensor instance trust
  - fusion system use/configuration
  - environment conditions
- Interested in a pedigree ontology that
  - can be extended to other types of data
  - has a connection to C2IEDM via Reporting-Data entity

Proposed Pedigree Ontology



Application Concept



Challenges

- Where does the pedigree data come from?
- How does it get tagged to legacy data?
  - suggestion: OTH-GOLD REMARK field
- Are the rewards of metadata collection worth the cost?

Recap

- Objectives
- Overview of OWL
- Pedigree as Metadata
- Proposed Pedigree Ontology
- A Candidate Application
- Challenges