A Probabilistic Ontology Development Methodology

International Command & Control Research & Technology Symposium

June 16-19, 2014

Richard J. Haberlin Jr.   *(EMSolutions)*

Paulo C.G. da Costa   *(George Mason University)*

Kathryn B. Laskey   *(George Mason University)*

EMSolutions

Background

An ontology is an explicit, formal representation of knowledge about a domain of application. This includes

– Types of entities that exist in the domain;

– Properties of those entities;

– Relationships among entities;

– Processes and events that happen with those entities;

where the term entity refers to any concept (real or fictitious, concrete or abstract) that can be described and reasoned about within the domain of application [Costa, 2005].

"An ontology is an explicit specification of a conceptualization [Gruber, 95].*"*

Ontology

- Ontologies provide a hierarchical structure of entity classes and a formal way of expressing their relationships
    - First-order expressivity
    - Supports logical reasoning
- There is significant literature on engineering traditional ontologies
- Ontologies lack built-in, principled support to adequately account for uncertainty
    - Annotating ontologies with simple probability annotations fails to convey structure of probabilistic representation
    - Less expressive probability schemes do not convey ontology structure, and so are inadequate

Probabilistic Ontology Defined

Ontology

A *probabilistic* ontology is an explicit, formal representation of knowledge about a domain of application. This includes

- – Types of entities that exist in the domain;
- – Properties of those entities;
- – Relationships among entities;
- – Processes and events that happen with those entities;

Uncertainty

- – **Statistical regularities that characterize the domain;**
- – **Inconclusive, ambiguous, incomplete, unreliable, and dissonant knowledge related to entities of the domain;**
- – **Uncertainty about all the above forms of knowledge;**

where the term entity refers to any concept (real or fictitious, concrete or abstract) that can be described and reasoned about within the domain of application [Costa, 2005].

A probabilistic ontology extends a traditional ontology to represent uncertainty.

- Integrates inferential reasoning power of probabilistic representations with first-order expressivity of ontologies

- Provides a means to represent and reason with uncertainty

- Limited literature on construction

Comprehensively describes knowledge about a domain and the uncertainty embedded in that knowledge in a principled, structured and sharable way [Brisset, 2003].

"It would be interesting to have a tool guiding the user on the steps necessary to create a probabilistic ontology and link this documentation to its implementation [Carvalho, 2011]."

EMSolutions

Why Probabilistic Ontologies?

- Suppose an ontology of organisms contains the following classes and relationships:

isA

isComposedOf
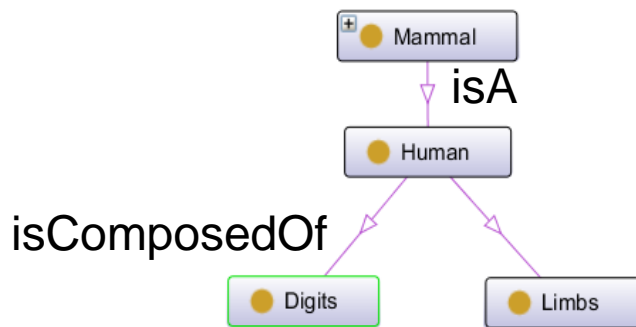
- Humans *usually* have:
  – 2 arms & 2 legs
  – 10 fingers & 10 toes
- However, if a man loses a limb….
  – Is he no longer human?

Premise of an argument can be uncertain (e.g. Humans have 2 legs): (in)validity of the argument imposes no condition on the certainty of the conclusion (an amputee is Human).

EMSolutions

- The Semantic Technologies (ST) community needed a comprehensive methodology for the development, implementation, and evaluation of probabilistic ontologies
  - Ontology use is on the rise
  - A means to incorporate uncertainty is a necessity
  - Limited literature on production of probabilistic ontologies
- Ontological engineering ensures ontologies developed for knowledge-sharing and reuse are explicit, logical and defensible
- Standard ontological engineering methods provide insufficient support for complexity of probabilistic ontology development

A similar methodology is needed for development of <u>probabilistic</u> ontologies

EMSolutions

Create a systematic approach to probabilistic ontology development

- Facilitated through a reference architecture
  - Formalizes the application of the methodology
  - Extensible to various domains
- Follows an iterative methodology applicable to any Systems Engineering development process
  - Allows continuous expansion and evaluation
  - Simplifies development and logic checking through spiraling
- Ensures the implemented design meets requirements

EMSolutions

The Process of Probabilistic Ontology Development

Reference Architecture for PO Development
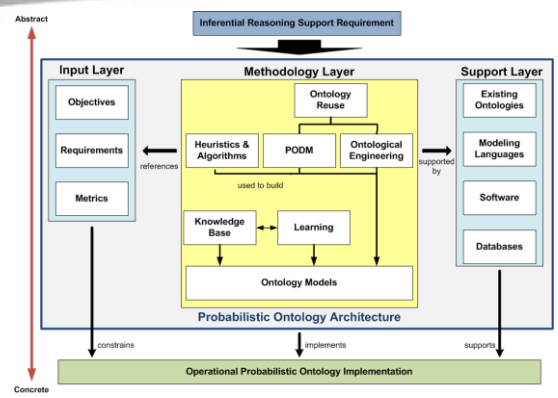
PO Development Methodology (PODM)
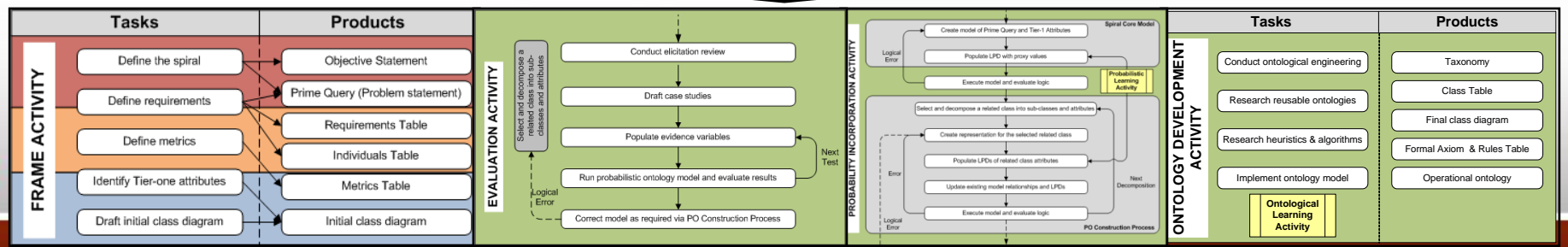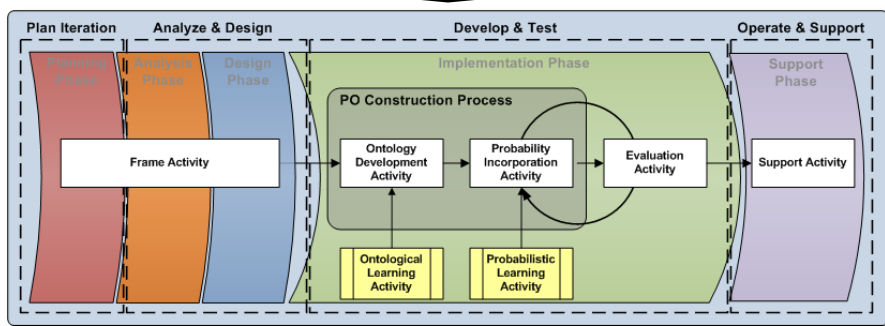
Probabilistic Ontology Development Methodology (PODM)

- PODM addresses evolution of requirements into an ontology that is probabilistically-integrated
  - Explicitly describes the iterative tasks required to produce a PO with in-situ evaluation steps
- Suitable for both spiral and waterfall development processes
- Application of PODM
  - Specific decision problem
  - Grounded in an inclusive ontology representing its entities
  - Incorporates probabilities to represent uncertainty

Establishes a solution grounded in an inclusive ontology representing its entities and incorporation of probabilities to represent uncertainty

EMSolutions

PODM in Spiral Development Cycle

Plan Next Spiral
(UP: Inception)

Analyze and Design
(UP: Elaboration)

Project Planning

Analysis Phase → Design Phase

SDLC Phase			
PODM Activity	**Waterfall**	**Spiral**	**Unified Process**
Frame	Planning Analysis Design	Plan Analyze & Design	Inception Elaboration
Ontology Development **Probability Incorporation** **Refinement** **Evaluation**	Implementation	Develop & Test	Construction
Operation	Support	Operate & Support	Transition

- ements
- s
- attributes
- class diagram

- Maintain system
- Upgrade system
- Support users

- Ontology development
- Probability incorporation
- Refinement
- Evaluation

Operate and Support
(UP: Transition)

Develop and Test Model
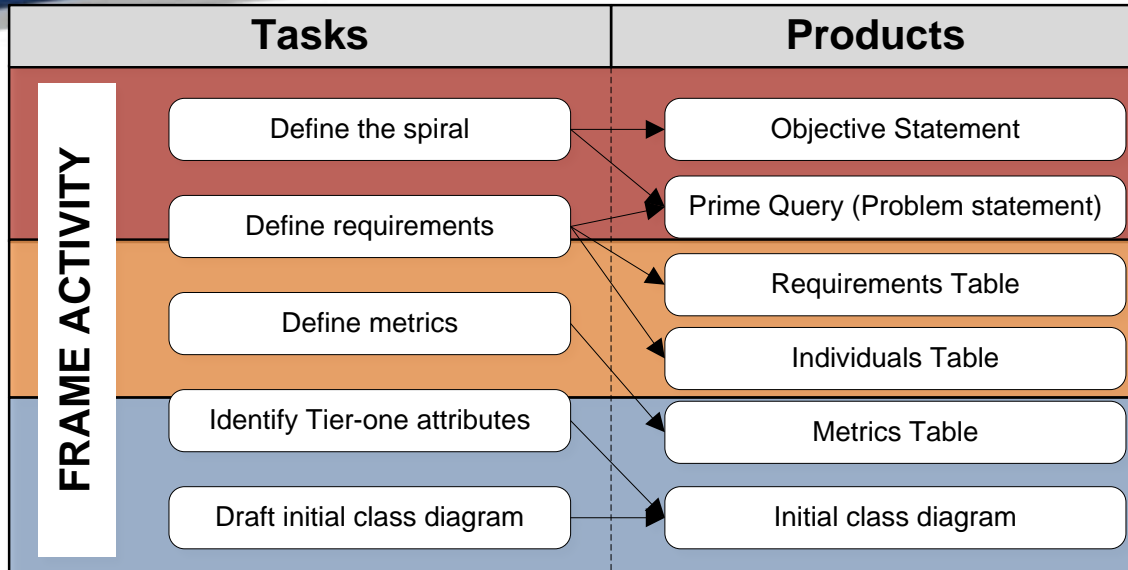(UP: Construction)

EMSolutions
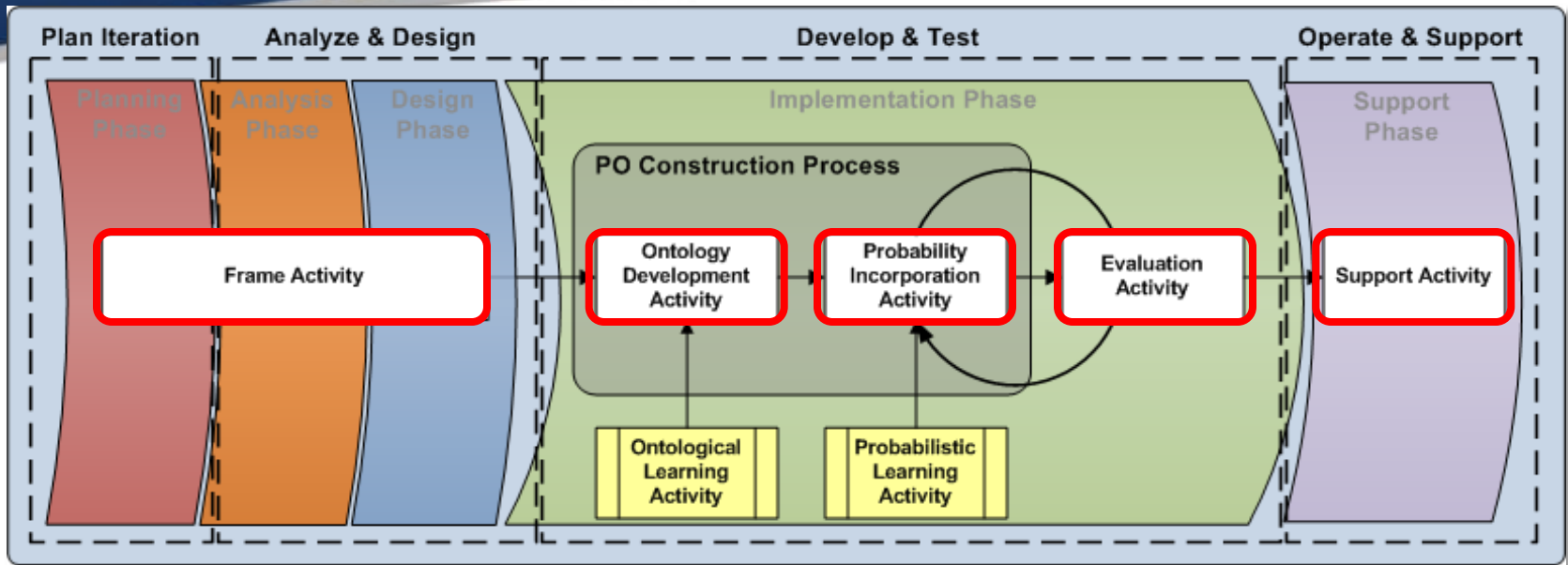
Maintenance, improvement and operational support tasks to maintain currency on the current build and support user operation.

processes
– Ensures interim steps evaluated for valid relationships and correct logic

Tasks	Products
FRAME ACTIVITY	
Define the spiral	Objective Statement
Define requirements	Prime Query (Problem statement)
Define metrics	Requirements Table
	Individuals Table
Identify Tier-one attributes	Metrics Table
Draft initial class diagram	Initial class diagram

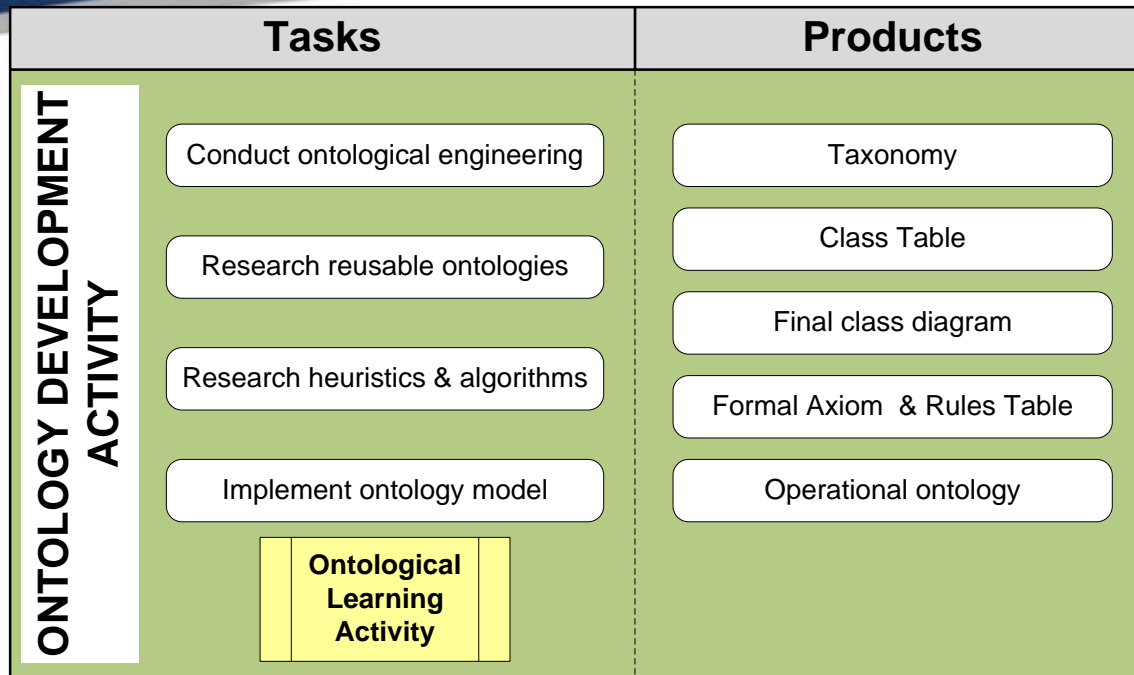- Based on the Overarching and Spiral Objective Statements
- Defines the Spiral Core Model
  - Prime Queries
  - Tier-one Attributes
- Scopes the spiral with the stakeholder's requirements and constraints

The Prime Queries and their associated Tier-one attributes define the spiral Core Model iterated in the construction cycle to create the necessary PO for inferential reasoning

EMSolutions

Ontology Development Activity

Tasks	Products
ONTOLOGY DEVELOPMENT ACTIVITY	
Conduct ontological engineering	Taxonomy
Research reusable ontologies	Class Table
Research heuristics & algorithms	Final class diagram
Implement ontology model	Formal Axiom & Rules Table
Ontological Learning Activity	Operational ontology

- Summarizes engineering tasks required to produce a working ontology
- Selection of an ontological engineering methodology is context dependent
- Fidelity of the ontological model is context dependent
- There are tasks and products common to each of these processes

EMSolutions

PROBABILITY INCORPORATION ACTIVITY

Spiral Core Model

- Create model of Prime Query and Tier-1 Attributes
- Populate LPD with proxy values
- Execute model and evaluate logic

Logical Error

Probabilistic Learning Activity

PO Construction Process

- Select and decompose a related class into sub-classes and attributes
- Create representation for the selected related class
- Populate LPDs of related class attributes
- Update existing model relationships and LPDs
- Execute model and evaluate logic

Error

Logical Error

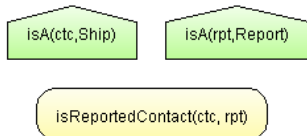Next Decomposition

Prime Query + Tier-1 Attributes = Spiral Core Model

Completed Military Ship PO
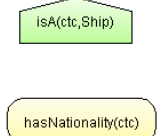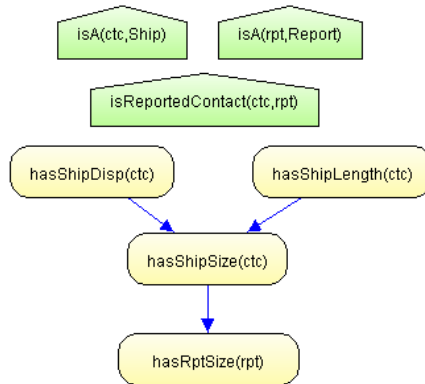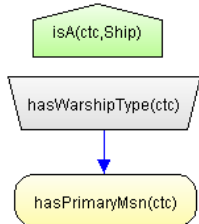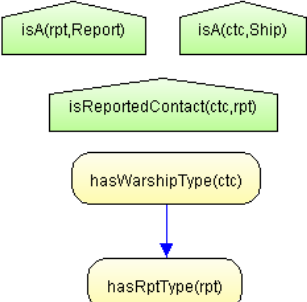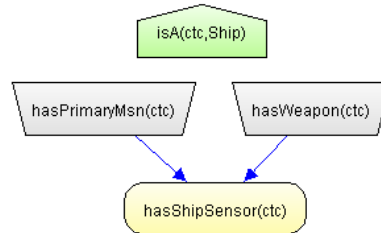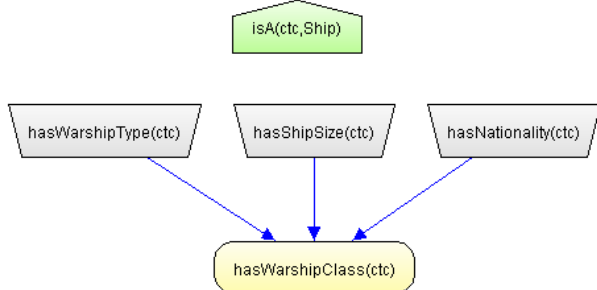
From an AOR-specific library (ontology), the MilShip PO infers the warship cass of an unknown contact based on limited or conflicting reports of varying pedigrees (uncertainty) .

Evaluation Activity

EVALUATION ACTIVITY

Select and decompose a related class into sub-classes and attributes

Conduct elicitation review

Draft case studies

R...

Logical Error

Co...
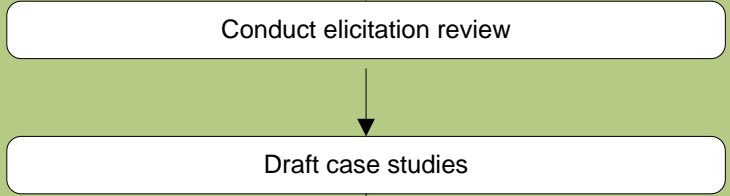
Evidence Table Case 4

Variable	Evidence
rs1	hasRptSize(rs1(Report))=Rpt_SizeMedium
ctc1, rs1	isReportedContact(ctc1(Ship), rs1(Report))=true
rt1	hasReportedType(rt1(Report))=Rpt_TypeFFG
ctc1, rt1	isReportedContact(ctc1(Ship), rt1(Report))=true
ctc1	hasSensor(ctc1(Ship))=RAS_SPY1D
ctc1	hasNationality(ctc1((Ship))=Nation_DE

hasRptSize__rs1
Rpt_SizeLarge	0%
Rpt_SizeMedium	100%
Rpt_SizeSmall	0%
absurd	0%

hasShipDisp__ctc1
Disp5kto10k	41.15%
DispLess5k	53.86%
DispGreater10k	4.99%
absurd	0%

hasShipSize__ctc1
SizeMedium	73.58%
SizeLarge	4.04%
SizeSmall	22.38%
absurd	0%

hasShipLength__ctc1
LengthGreater150	4.99%
Length125to150	41.15%
LengthLess125	53.86%
absurd	0%

hasWarshipClass__ctc1
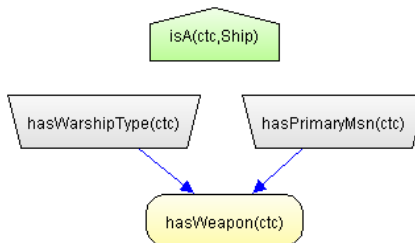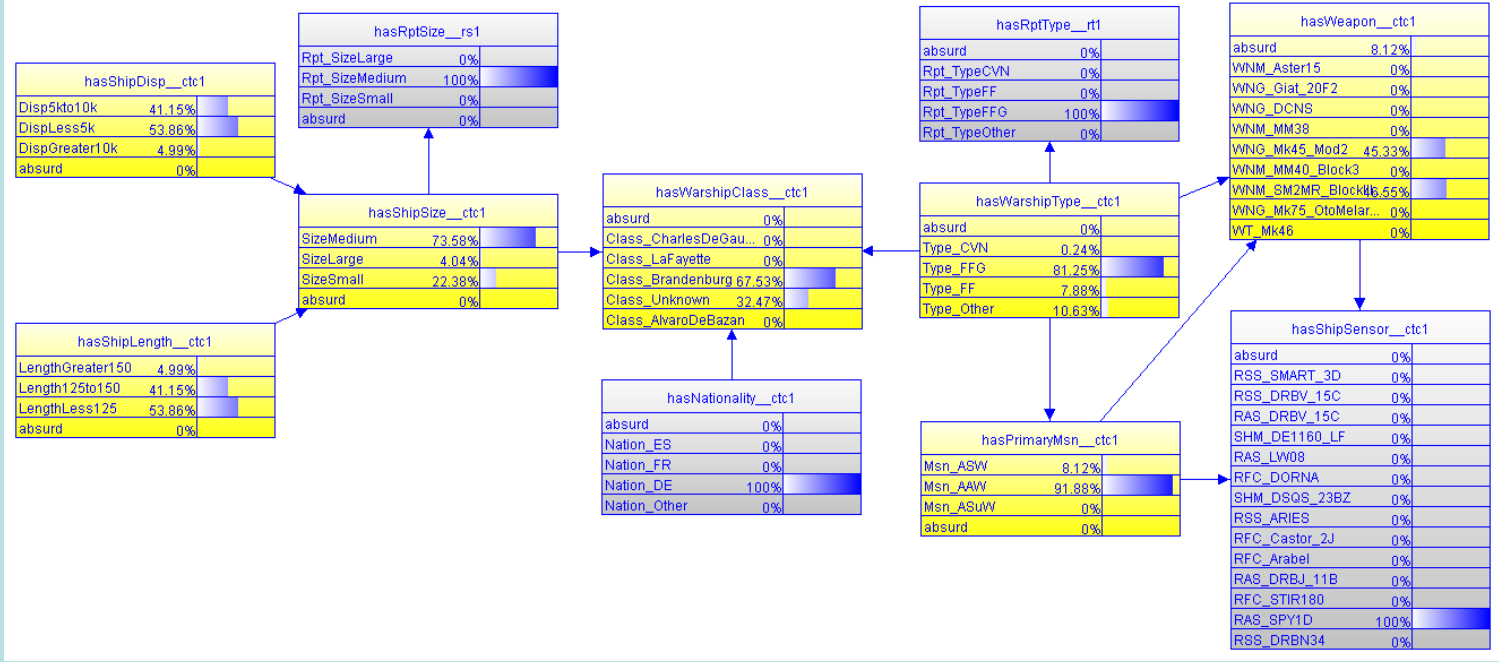|	
absurd	0%
Class_CharlesDeGau...	0%
Class_LaFayette	0%
Class_Brandenburg	67.53%
Class_Unknown	32.47%
Class_AlvaroDeBazan	0%

hasNationality__ctc1
absurd	0%
Nation_ES	0%
Nation_FR	0%
Nation_DE	100%
Nation_Other	0%

hasRptType__rt1
absurd	0%
Rpt_TypeCVN	0%
Rpt_TypeFF	0%
Rpt_TypeFFG	100%
Rpt_TypeOther	0%

hasWarshipType__ctc1
absurd	0%
Type_CVN	0.24%
Type_FFG	81.25%
Type_FF	7.88%
Type_Other	10.63%

hasPrimaryMsn__ctc1
Msn_ASW	8.12%
Msn_AAW	91.88%
Msn_ASuW	0%
absurd	0%

hasWeapon__ctc1
absurd	8.12%
WNM_Aster15	0%
WNG_Giat_20F2	0%
WNG_DCNS	0%
WNM_MM38	0%
WNG_Mk45_Mod2	45.33%
WNM_MM40_Block3	0%
WNM_SM2MR_Block...	46.55%
WNG_Mk75_OtoMelar...	0%
WT_Mk46	0%

hasShipSensor__ctc1
absurd	0%
RSS_SMART_3D	0%
RSS_DRBV_15C	0%
RAS_DRBV_15C	0%
SHM_DE1160_LF	0%
RAS_LW08	0%
RFC_DORNA	0%
SHM_DSQS_23BZ	0%
RSS_ARIES	0%
RFC_Castor_2J	0%
RFC_Arabel	0%
RAS_DRBJ_11B	0%
RFC_STIR180	0%
RAS_SPY1D	100%
RSS_DRBN34	0%

- Elicitation
- Case stud...
  - Cases
    - Ty...
  - Errone...

EMSolutions

Test cases ensure model meets requirements specified by objective

Problem

- Ontological engineering methodologies are unsuitable for production of probabilistic ontologies.

- The literature on probabilistic ontology development is extremely limited.

Solution

- Reference Architecture for Probabilistic Ontology Development

- Probabilistic Ontology Development Methodology

EMSolutions

A Probabilistic Ontology Development Methodology

Richard Haberlin
Paulo Costa
Kathy Laskey

GEORGE MASON UNIVERSITY

EMSolutions