Cross-Domain Ontology Resolution in Net-Centric Command and Control

Dale E. Lichtblau, Steven P. Wartik
Institute for Defense Analyses

ICCRTS 2007

Outline

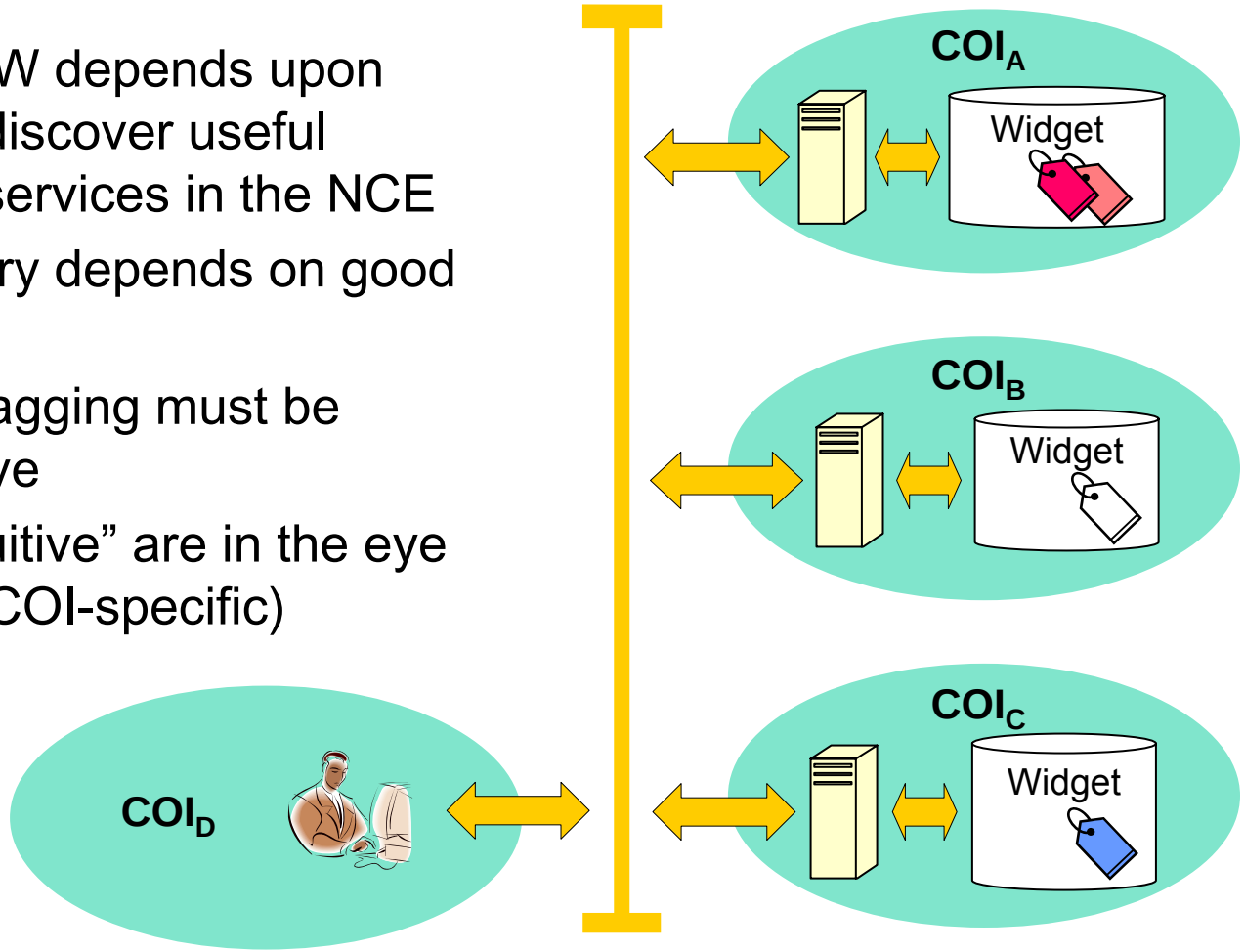
- **Motivation**
  - Importance of discovery and tagging in a net-centric environment
- **Problem statement**
  - Cross-ontology term resolution
- **The LOM approach to automated term resolution**
  - Li's approach and its limitations
- **Domain-specific knowledge extension**
  - A proposed improvement
- **Summary**

Motivation

Importance of Discovery and Tagging in the NCE

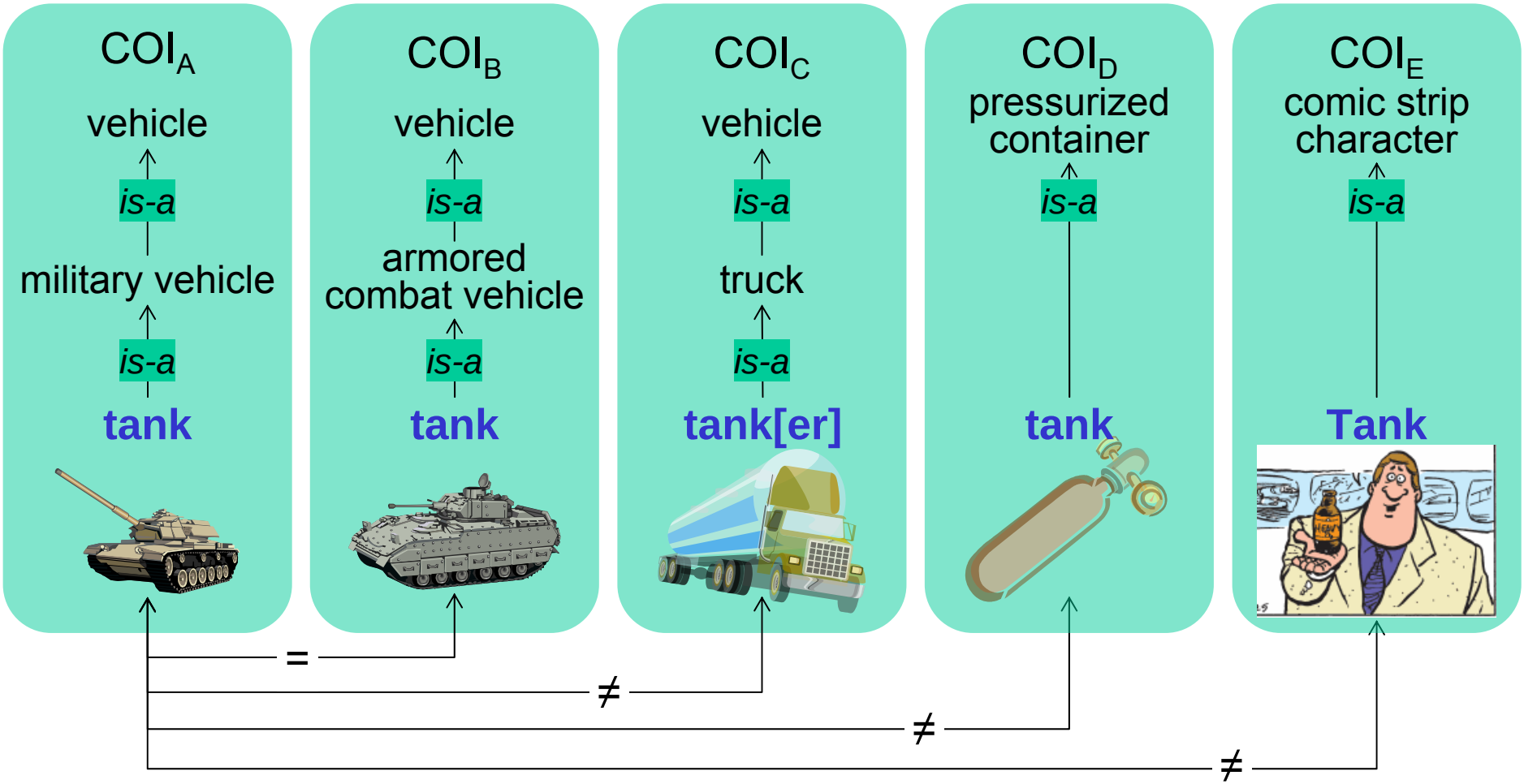
Assertions

1. Success of NCOW depends upon ability to readily discover useful information and services in the NCE
2. Effective discovery depends on good semantic tagging
3. Good semantic tagging must be sound and intuitive
4. “Sound” and “intuitive” are in the eye of the beholder (COI-specific)



Problem Statement

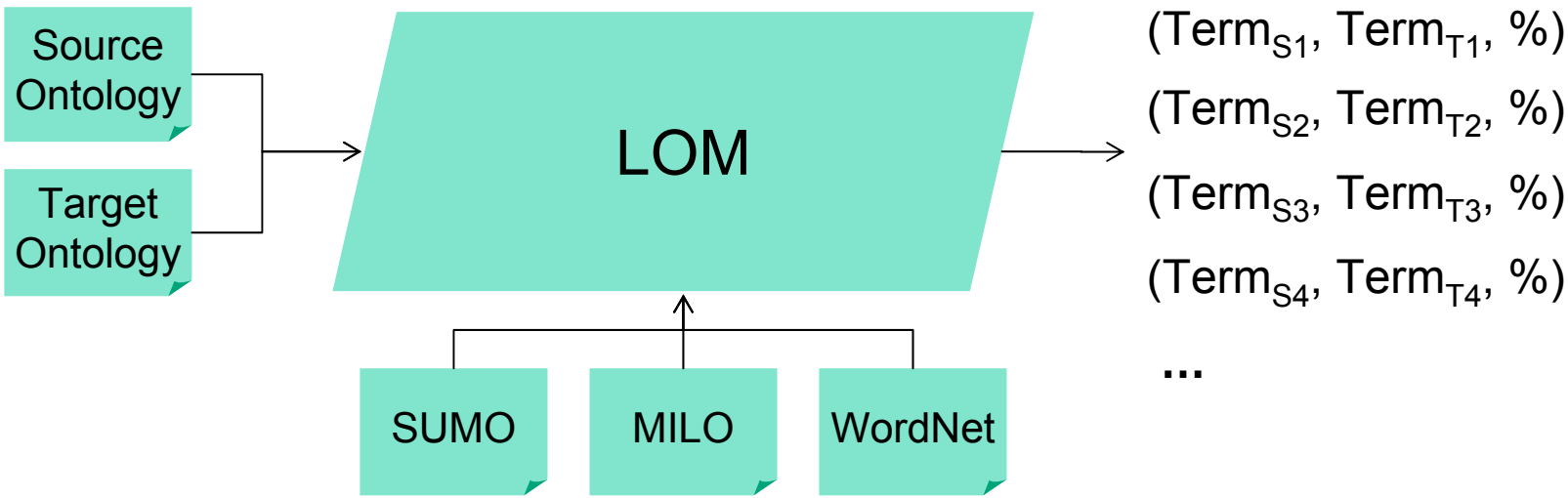
Cross-Ontology Term Resolution



- Each COI maintains an ontology
- How do we decide which terms are equivalent?

The LOM Approach

- LOM: the Lexicon-based Ontology Mapping tool
- Developed by John Li at Teknowledge Corporation
- 71% precision, 57% recall



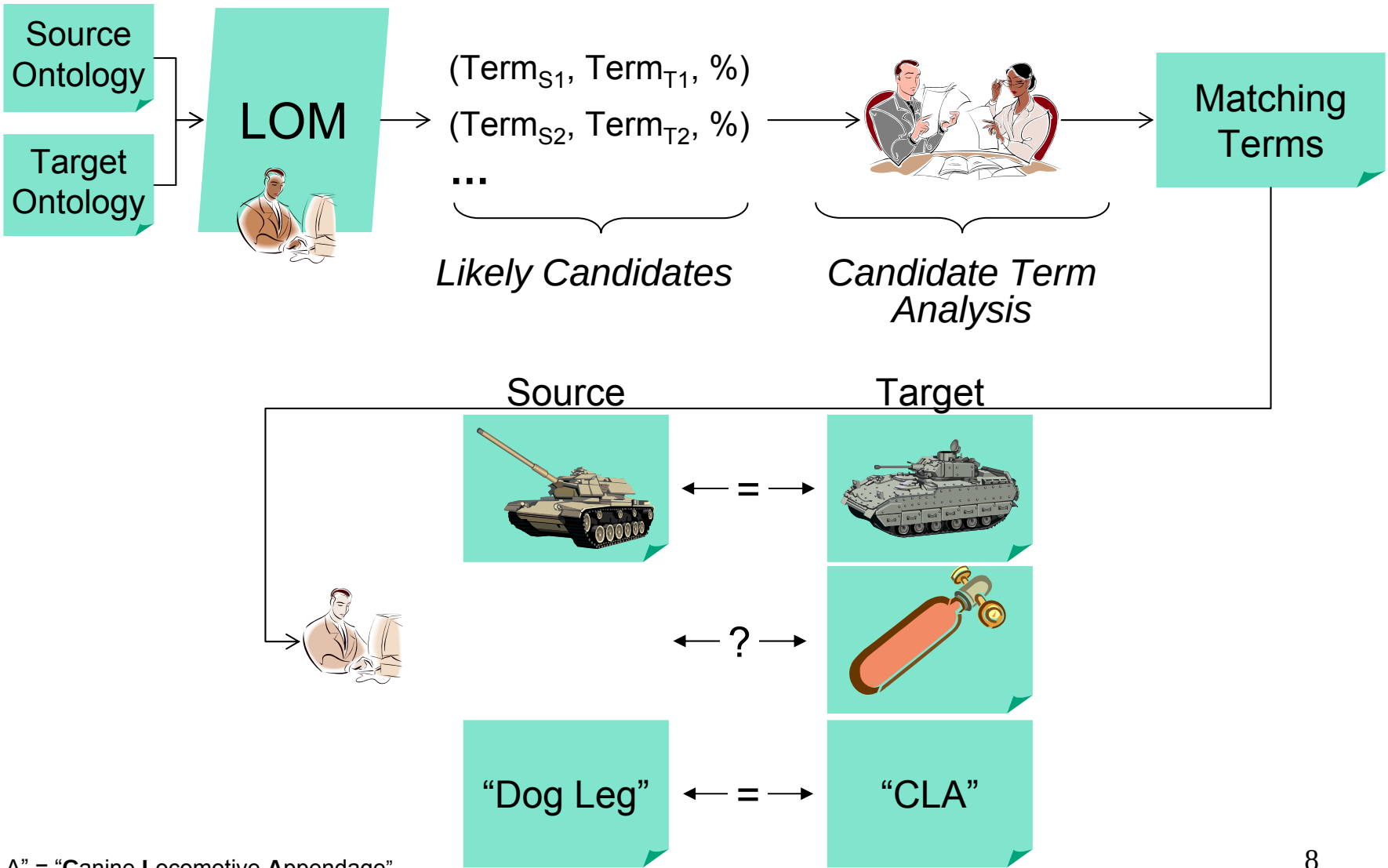
The LOM Algorithm

1. Match whole terms
   - Look for identically-named terms in source and target ontologies
   - Ignore case distinctions (but nothing else)
2. Match word constituents
   - Identify compound terms (“ObjectItem” becomes “Object Item”)
   - Consider punctuation, stemming
   - Filter common terms with stop list
3. Match synonym sets
   - Use WordNet to find synonyms
4. Match types
   - Use SUMO and MILO to refine synonym sets

How LOM Assigns Confidence Ratings

- Step 1 (match whole terms) assigns 0/1 rating
- Subsequent steps assign percentage based on ratio between number of matched words and average number of words being considered
- Each step has an empirically-derived weight factor
- Final score for a matching pair at each step is maximum of its score from individual step times weight factor
- Minimum level filters out unlikely matches

LOM Process for Cross-Ontology Term Resolution



"CLA" = "**C**anine **L**ocomotive **A**ppendage"

Strengths and Limitations of Li's Algorithm

- General-purpose approach
- Incorporates no domain-specific knowledge
- Uses SUMO (upper-level) and MILO (mid-level)

Can a domain-specific approach do any better?

Assumptions

- Ontologies are expressed in Web Ontology Language (OWL)
- C2 domain-specific terminology is drawn from Joint Command, Control, and Consultation Information Exchange Data Model (JC3IEDM)

A C2-Specific Algorithm for Cross-Ontology Term Resolution

1. Match whole terms
   - 1a. Match term IDs
   - 1b. Match term labels
   - 1c. Match terms in element definitions, using OWL comments
2. Match word constituents
3. Match synonym sets
   - 3a. Perform property comparisons
   - 3b. Classify the ontologies
4. Map terms to JC3IEDM terms

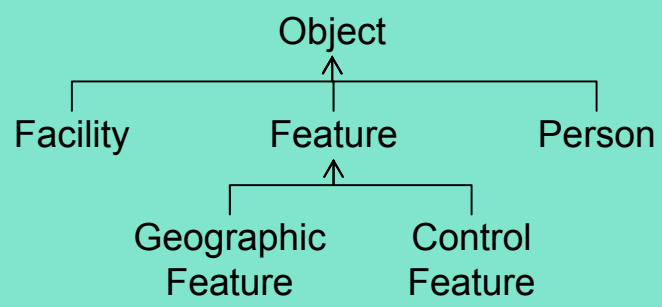
Property Comparisons (Step 3a)

- OWL properties can
  - Be functional
  - Be inverse-functional
  - Define their domain/range
  - Define cardinality

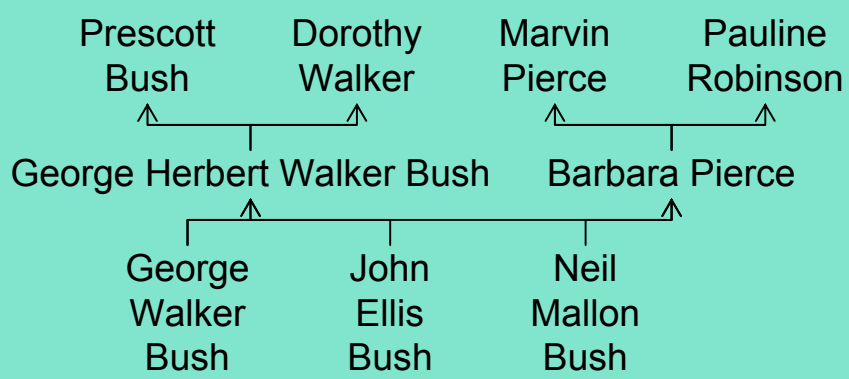
Y
↑ = "X has parent Y"
X

But "has-parent" is not the same property!

A Functional Hierarchy



A Genealogical Hierarchy



Classifying Ontologies (Step 3b)

- OWL supports description-logic reasoning
- Identification of a *contradiction* (a class with no members) shows a mistaken assumption
- We capitalize on this as follows
  - Steps 1 through 3a constitute a set of assumptions about term equivalence
  - We execute a reasoner using the ontologies we're comparing and the assumptions about terms
  - Each contradiction reveals an incorrect assumption about term equivalence

Map terms to JC3IEDM terms (Step 4)

- Similar to LOM's step 4, but we use JC3IEDM instead of SUMO and MILO
- Goal: test equivalence of an OWL class to a JC3IEDM concept represented as an OWL class
- JC3IEDM classes are disjoint—therefore two classes are only equivalent if they're subclasses of the same class

Summary

- NCE is intended to be an environment in which automated agents can discover information and services
- A prerequisite of discovery is the ability to infer the “meaning” of a discovered term
- Ontologies have been proposed as one mechanism to enable this inference
- There must be a means to discover the conceptual similarities between ontologies
- LOM is a tool to create mappings between ontologies
- Our extension to LOM leverages the similarities between C2-specific ontologies that deal with command and control
- Use of the JC3IEDM provides a more specific context for term resolution than does the use of SUMO and MILO
- Like LOM, our approach requires some human intervention
- Technology has not reached the point where automated tools can be trusted to infer contextual similarity
- However, improvement in the capability to discover information and services is necessary to achieve the promise of net-centricity in C2