Exposing and Capturing Mapping Relationships

Mala Mehrotra
Pragati Synergetic Research Inc.

Ontolog Forum Panel – Rationale, Expectations & Requirements
3rd April 2008

Work supported by AFRL, ONR & DARPA

OOR needs for content /application providers

- *Content developers*: Discover related terms/axioms/models for reuse
  - Context – collaboration groups of concepts
    - *region* (geographic, biological, political)
  - Depth/detail
    - *month* in SUMO vs. *monthDescription* in DAML time ontologies
  - Differences in competing models
    - *TimeInterval* in SUMO vs *DurationDescription* in DAML
  - Degree of Crossover/Overlap
    - More than just imports closure
    - Orthogonality measures across ontologies
- *Application developers*: Interoperate using multiple ontologies
  - Create formalized mapping relationships
  - Find mapping relationships

Infrastructure Needs

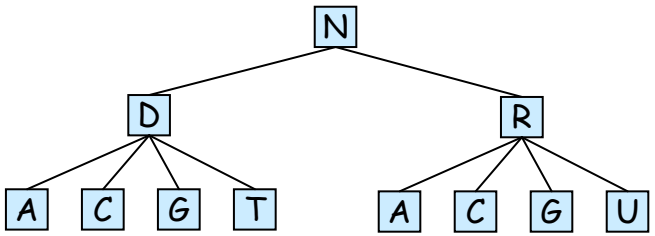
- Cognitive Tools for discovery
  - Collaborating groups of concepts used in applications
  - Implicit relationships across resources
  - Ontological/Taxonomy hierarchy browsing
  - Human-machine collaboration mode
- Mapping Tools for capturing inter-resources' relationships
- Need formal representation of relationships for reasoners
  - A large repertoire of relationships
  - Multiple ontological representations
  - Mechanisms to represent formalism in human-readable form

Exposing Shades of Relationships

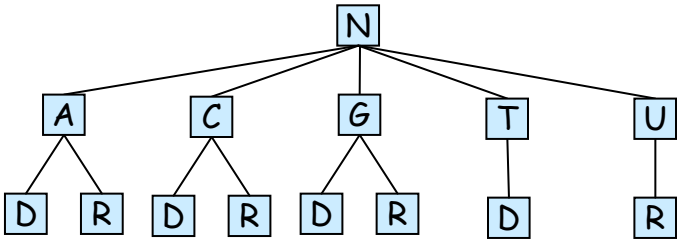
- Equivalence
  - *PhDThesis* & *DoctoralThesis*
- Partial
  - Specialization
  - Generalization
  - Restriction (various types) on properties
- Inverse (swapping of arguments, argument permutations)
  - *move-in* vs. *move-out*
- Negation
- Ternary
  - Transitive (multi-argument mappings)
    - *Task-unit* connected to *tasked-unit* through *country codes;*
  - Clichés (generalization of a repeating pattern)
    - Same type of initialization process over various types of instruments

Model equivalence despite terminology differences

- A nucleotide molecule, in Cyc's BiochemistryMt is represented by
  - holding the *sugars* constant at first level and varying the *base* (left figure) or
  - holding the *base* constant at first level and varying the *sugar* (right figure)
- The left representation good for chain type of reasoning for the molecule that is at the nucleotide level.
- The right representation good for the matching base pair type of level of reasoning.
- Clustering brought to attention both these representations.



Sugar-dependent representation



Base-dependent representation

Model equivalence despite terminology differences

Sugar-dependent representation

```
(#$genls #$Thymine-Deoxyribonucleotide  #$Deoxyribonucleotide)
(#$genls #$Adenine-Deoxyribonucleotide  #$Deoxyribonucleotide)
(#$genls #$Cytosine-Deoxyribonucleotide #$Deoxyribonucleotide)
(#$genls #$Guanine-Deoxyribonucleotide  #$Deoxyribonucleotide)
```

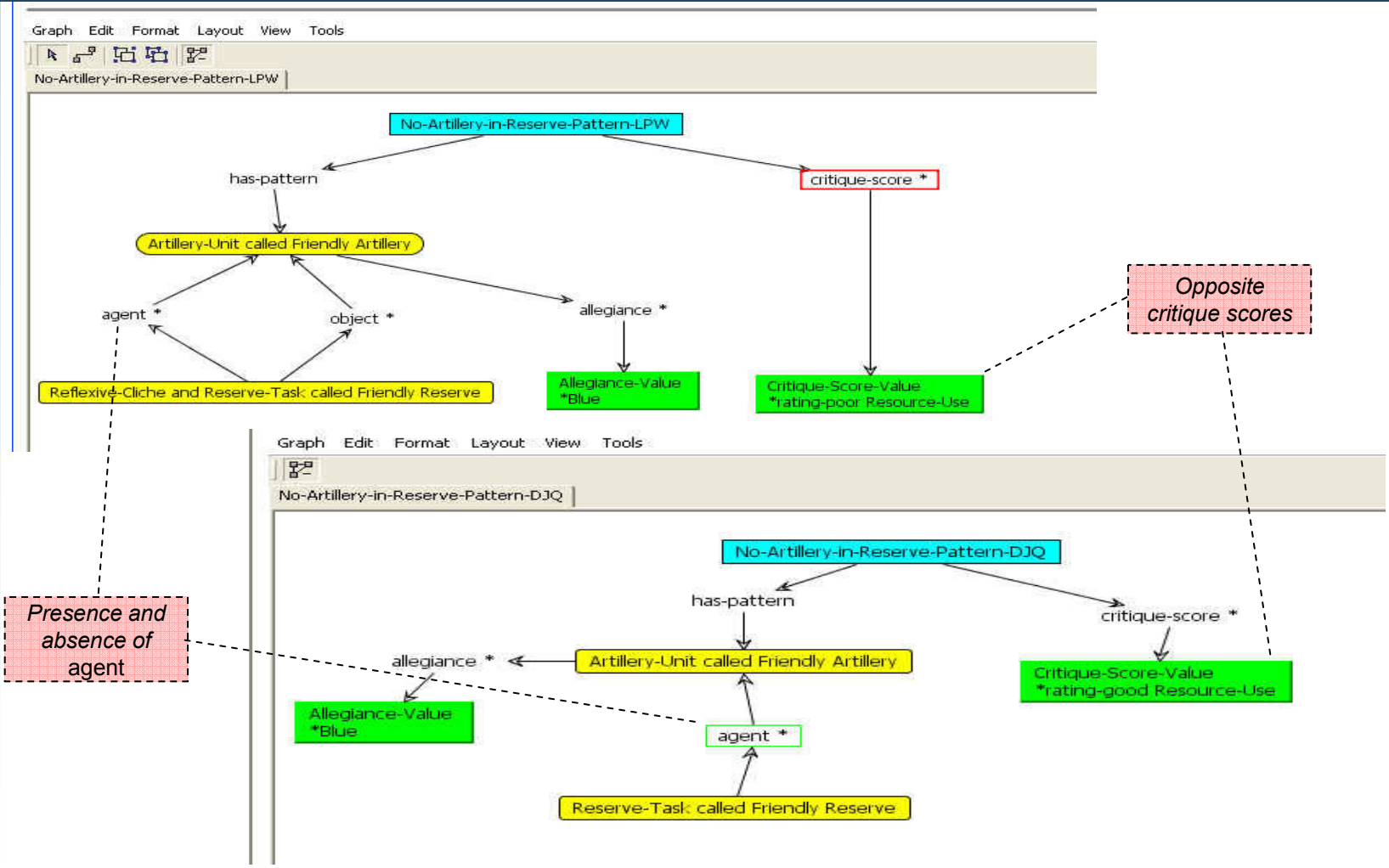
```
(#$genls #$Uracil-Ribonucleotide   #$Ribonucleotide)
(#$genls #$Adenine-Ribonucleotide  #$Ribonucleotide)
(#$genls #$Cytosine-Ribonucleotide #$Ribonucleotide)
(#$genls #$Guanine-Ribonucleotide  #$Ribonucleotide)
```

Axiom Clusters showing multiple legitimate representations of Nucleotides from Cyc's BioChemistryMt

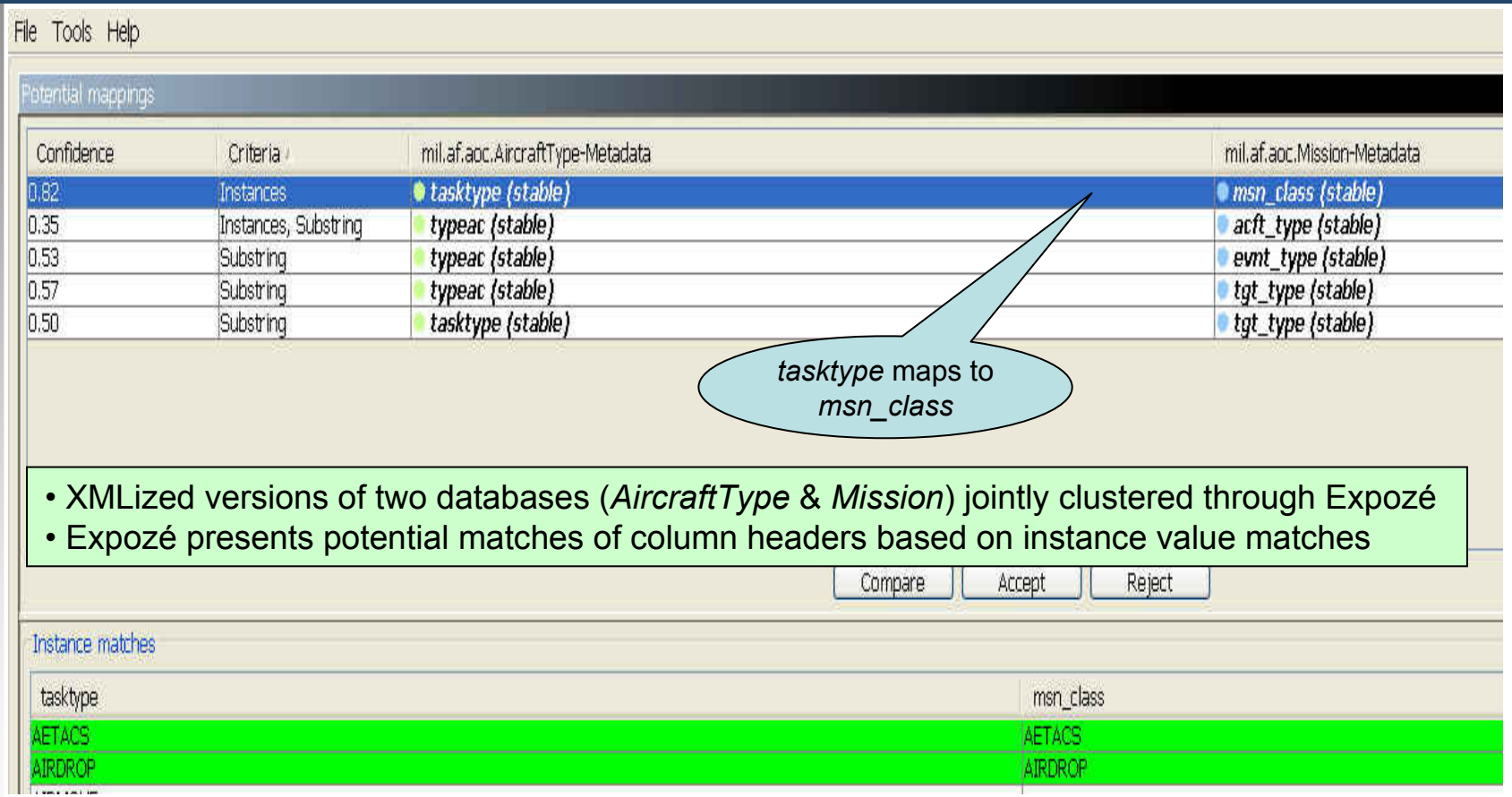
Base-dependent representation

```
(#$genls #$Nucleotide #$Molecule)
(#$genls #$Deoxyribonucleotide #$Nucleotide)
(#$genls #$Ribonucleotide #$Nucleotide)
(#$genls #$AdenineNucleotide #$Nucleotide)
(#$genls #$CytosineNucleotide #$Nucleotide)
(#$genls #$GuanineNucleotide #$Nucleotide)
(#$genls #$Adenine-Ribonucleotide #$AdenineNucleotide)
(#$genls #$Adenine-Deoxyribonucleotide #$AdenineNucleotide)
(#$genls #$Cytosine-Deoxyribonucleotide #$CytosineNucleotide)
(#$genls #$Cytosine-Ribonucleotide #$CytosineNucleotide)
(#$genls #$Guanine-Deoxyribonucleotide #$GuanineNucleotide)
(#$genls #$Guanine-Ribonucleotide #$GuanineNucleotide)
```

Equivalent concepts with Negation



Equivalence Mapping across Database Schemas in DISCOVER



Mapping Axioms in CL

(SubTable C *AircraftType*)

(Domain (FromTo *tasktype msn_class*) *AircraftType*)

(Range (FromTo *tasktype msn_class*) *Mission*)

```
(forall ((x C))
   (= (tasktype x)(msn_class
         ((FromTo tasktype msn_class) x) )))
```

CL formulas have been developed in collaboration with Dr. Pat Hayes (IHMC)

Partial Mapping across Database Schemas in DISCOVER

Potential mappings

Confidence	Criteria	mil.af.aoc.FriendlyBase-Metadata	mil.af.aoc.Mission-Metadata
1.00	Exact	*comments (stable)*	*comments (stable)*
1.00	Exact	*update_date_time (stable)*	*update_date_time (stable)*
0.13	Instances	*closedatetime (stable)*	*start_time (stable)*
0.04	Instances	*closedatetime (stable)*	*end_time (stable)*
0.06	Instances	*baseid (stable)*	*evnt_loc (stable)*
0.23	Instances	*untildatetime (stable)*	*start_time (stable)*
0.14	Instances	*untildatetime (stable)*	*end_time (stable)*
0.02	Instances	*effectivedatetime (stable)*	*update_date_time (stable)*
0.46	Instances	*effectivedatetime (stable)*	*start_time (stable)*
0.26	Instances	*effectivedatetime (stable)*	*end_time (stable)*
0.15	Instances	*opendatetime (stable)*	*start_time (stable)*
0.12	Instances	*opendatetime (stable)*	*end_time (stable)*
0.29	Instances	*update_date_time (stable)*	*start_time (stable)*

Compare | Accept | Reject

Instance matches

baseid	evnt_loc
KDPG	KDPG
KDUW	KDUW
KEDW	
KEND	
KEYL	KEYL
KFHU	KFHU
KFMH	
KGSB	
KGTF	
KHMN	
KIAB	

evnt_loc maps to *baseid*
if *eventType* is *take-off* or *landing*

Mapping Axioms in CL

- Mapping of *baseid* (FriendlyBase) to *evnt_loc (*Mission) only during aircraft *Takeoff* or *Landing*
- Axioms assume that this information is recorded in another table called *Event* with a column *eventType*
- Assumption allows us to define the subtable *C1* by using a *FromTo* mapping from *baseid* to *eventType*
- Mapping of *baseid* (FriendlyBase) to *evnt_loc (*Mission) can then be defined as before using the subtable *C1*

(SubTable C1 *FriendlyBase*)

```
(forall (x)(iff
  (C1 x)
  (or (= (eventType ((FromTo baseid eventType) x) Landing))
      (= (eventType ((FromTo baseid eventType) x) Takeoff))       )))
```

```
(forall ((x C1))
   (= (baseid x)(evnt_loc
                ((FromTo baseid evnt_loc) x) )))
```

Conclusions

- Mapping relationships useful in:
  - Federated Query
  - Reuse/inspect
  - Interoperate
  - Fuse/integrate
  - Tracking Content (set triggers)
- Discovery of relationships requires cognitive aid tools
  - Human collaboration essential
- Applications will need formalized representation of the mappings in OOR

Contact Information

Contact Information:

Mala Mehrotra

Email: mm@pragati-inc.com
Voice: (650) 625-0274

Pragati Synergetic Research, Inc.
NASA Research Park Suite 2001
MS 19-46Q
NASA Ames Research Center
Moffett Field CA 94035
www.pragati-inc.com