

EULER/X: A Toolkit for Logic-based Taxonomy Integration

Mingmin Chen¹, Shizhuo Yu¹, Nico Franz²,
Shawn Bowers³ and Bertram Ludäscher¹

¹Dept. of Computer Science, UC Davis

²School of Life Sciences, Arizona State University

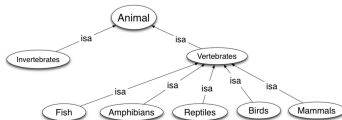
³Dept. of Computer Science, Gonzaga University

22nd International Workshop
on Functional and (Constraint) Logic Programming
Kiel, Germany

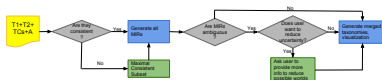
Sep 12, 2013

Outline

► Introduction



► EULER/X Workflow



► Summary

Challenge — describing classification provenance beyond synonymy

⇒ *Andropogon spp.* in the Carolinas, from Hackel 1889 to Weakley 2005

| Weakley 2005 | C. Campbell (1983, FNA 2003) | Godfrey & Wooten 1979 | RAB 1968 | Hitchcock & Chase 1950 | Blomquist 1948 | Small 1933 | Hackel 1889 |
|---|---|--|----------------------|---|--|----------------------|---|
| <i>Andropogon capillipes</i> var. <i>capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> "drylands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>glaucus</i> |
| <i>Andropogon capillipes</i> var. <i>dealbatus</i> (in prep.) | <i>A. virginicus</i> var. <i>glaucus</i> "wetlands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>dealbatus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "old-field variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "smooth variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon glaucopsis</i> | <i>A. glomeratus</i> var. <i>glaucopsis</i> | <i>A. glaucopsis</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>glaucopsis</i> |
| <i>Andropogon glomeratus</i> var. <i>hirsutior</i> | <i>A. glomeratus</i> var. <i>hirsutior</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>hirsutior</i> | ? | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>hirsutior</i> |
| <i>Andropogon glomeratus</i> var. <i>glomeratus</i> | <i>A. glomeratus</i> var. <i>glomeratus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>abbreviatus</i> |
| <i>Andropogon tenuispatheus</i> | <i>A. glomeratus</i> var. <i>pumilus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. virginicus</i> var. <i>tenuispatheus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>genuinus</i> |
| 5 species, 8 vars. | 2 species, 7 vars (+ 2 informal "variants") | 3 species, 4 vars. | 1 species | 3 species, 5 vars. | 3 species, 5 vars. | 3 species, 3 vars. | 2 species, 7 vars. |

Source: Weakley. 2005. Flora of the Carolinas, Virginia, and Georgia. Available at <http://www.herbarium.unc.edu/flora.htm>

Challenge — describing classification provenance beyond synonymy

⇒ *Andropogon spp.* in the Carolinas, from Hackel 1889 to Weakley 2005

| Weakley 2005 | C. Campbell (1983, FNA 2003) | Godfrey & Wooten 1979 | RAB 1968 | Hitchcock & Chase 1950 | Blomquist 1948 | Small 1933 | Hackel 1889 |
|---|---|--|----------------------|---|--|----------------------|---|
| <i>Andropogon capillipes</i> var. <i>capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> "drylands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>glaucus</i> |
| <i>Andropogon capillipes</i> var. <i>dealbatus</i> (in prep.) | <i>A. virginicus</i> var. <i>glaucus</i> "wetlands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>dealbatus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "old-field variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "smooth variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon glaucopsis</i> | <i>A. glomeratus</i> var. <i>glaucopsis</i> | <i>A. glaucopsis</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>glaucopsis</i> |
| <i>Andropogon glomeratus</i> var. <i>hirsutior</i> | <i>A. glomeratus</i> var. <i>hirsutior</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>hirsutior</i> | ? | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>hirsutior</i> |
| <i>Andropogon glomeratus</i> var. <i>glomeratus</i> | <i>A. glomeratus</i> var. <i>glomeratus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>abbreviatus</i> |
| <i>Andropogon tenuispatheus</i> | <i>A. glomeratus</i> var. <i>pumilus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. virginicus</i> var. <i>tenuispatheus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>genuinus</i> |
| 5 species, 8 vars. | 2 species, 7 vars (2 informal "variants") | 3 species, 4 vars. | species | 3 species, 5 vars. | 3 species, 5 vars. | 3 species, 3 vars. | 2 species, 7 vars. |

⇒ Individual columns represent past classifications of *Andropogon*.

Challenge — describing classification provenance beyond synonymy

⇒ *Andropogon spp.* in the Carolinas, from Hackel 1889 to Weakley 2005

| Weakley 2005 | C. Campbell (1983, FNA 2003) | Godfrey & Wooten 1979 | RAB 1968 | Hitchcock & Chase 1950 | Blomquist 1948 | Small 1933 | Hackel 1889 |
|---|---|--|----------------------|---|--|----------------------|---|
| <i>Andropogon capillipes</i> var. <i>capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> "drylands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>glaucus</i> |
| <i>Andropogon capillipes</i> var. <i>dealbatus</i> (in prep.) | <i>A. virginicus</i> var. <i>glaucus</i> "wetlands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>dealbatus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "old-field variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "smooth variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon glaucopsis</i> | <i>A. glomeratus</i> var. <i>glaucopsis</i> | <i>A. glaucopsis</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>glaucopsis</i> |
| <i>Andropogon glomeratus</i> var. <i>hirsutior</i> | <i>A. glomeratus</i> var. <i>hirsutior</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>hirsutior</i> | ? | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>hirsutior</i> |
| <i>Andropogon glomeratus</i> var. <i>glomeratus</i> | <i>A. glomeratus</i> var. <i>glomeratus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>abbreviatus</i> |
| <i>Andropogon tenuispatheus</i> | <i>A. glomeratus</i> var. <i>pumilus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. virginicus</i> var. <i>tenuispatheus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>genuinus</i> |
| 6 species, 8 vars. | 2 species, 7 vars (+ 2 informal "variants") | 0 species, 4 vars. | 1 species | 0 species, 6 vars. | 0 species, 6 vars. | 0 species, 0 vars. | 2 species, 7 vars. |

⇒ Individual rows represent equivalent taxonomic entities, (almost) regardless of their name labels.

Challenge — describing classification provenance beyond synonymy

⇒ *Andropogon spp.* in the Carolinas, from Hackel 1889 to Weakley 2005

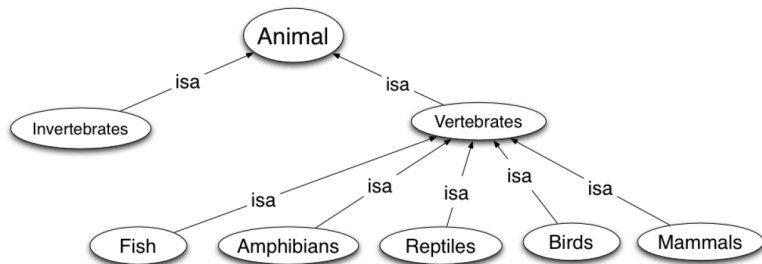
| Weakley 2005 | C. Campbell (1983, FNA 2003) | Godfrey & Wooten 1979 | RAB 1968 | Hitchcock & Chase 1950 | Blomquist 1948 | Small 1933 | Hackel 1889 |
|---|---|--|----------------------|---|--|----------------------|---|
| <i>Andropogon capillipes</i> var. <i>capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> "drylands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>glaucus</i> |
| <i>Andropogon capillipes</i> var. <i>dealbatus</i> (in prep.) | <i>A. virginicus</i> var. <i>glaucus</i> "wetlands variant" | <i>A. capillipes</i> | <i>A. virginicus</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glaucus</i> subvar. <i>dealbatus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "old-field variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> "smooth variant" | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>decipiens</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> var. <i>virginicus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>Andropogon glaucopsis</i> | <i>A. glomeratus</i> var. <i>glaucopsis</i> | <i>A. glaucopsis</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. virginicus</i> var. <i>glaucopsis</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>glaucopsis</i> |
| <i>Andropogon glomeratus</i> var. <i>hirsutior</i> | <i>A. glomeratus</i> var. <i>hirsutior</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>hirsutior</i> | ? | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>hirsutior</i> |
| <i>Andropogon glomeratus</i> var. <i>glomeratus</i> | <i>A. glomeratus</i> var. <i>glomeratus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>abbreviatus</i> |
| <i>Andropogon tenuispatheus</i> | <i>A. glomeratus</i> var. <i>pumilus</i> | <i>A. virginicus</i> var. <i>abbreviatus</i> | <i>A. virginicus</i> | <i>A. glomeratus</i> | <i>A. virginicus</i> var. <i>tenuispatheus</i> | <i>A. glomeratus</i> | <i>A. macrochrous</i> var. <i>genuinus</i> |
| 6 species, 8 vars. | 2 species, 7 vars (+ 2 informal "variants") | 0 species, 4 vars. | 1 species | 0 species, 6 vars. | 0 species, 6 vars. | 0 species, 0 vars. | 0 species, 7 vars. |

⇒ Individual rows represent equivalent taxonomic entities, (almost) regardless of their name labels.

⇒ Name/synonymy relationships are not sufficiently granular to capture this evolution of taxonomic views of *Andropogon* species.

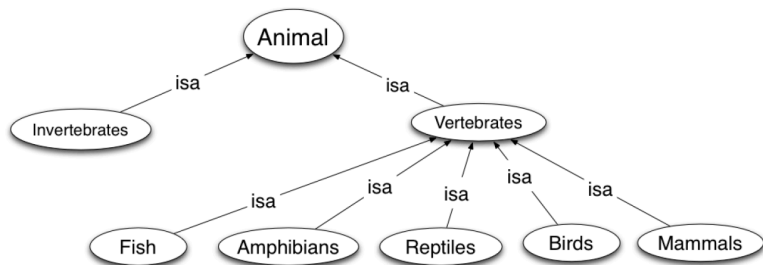
Taxonomy & Taxonomic Constraints

A **taxonomy** is a set of biological taxa (concepts) in an *ISA* hierarchy with *taxonomic constraints*:



Taxonomy & Taxonomic Constraints

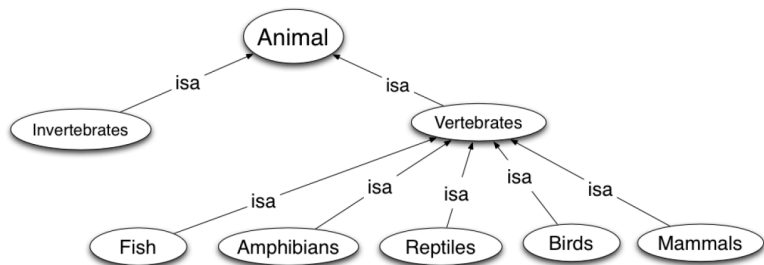
A **taxonomy** is a set of biological taxa (concepts) in an *ISA* hierarchy with *taxonomic constraints*:



- ▶ Non-Emptiness: Birds $\neq \emptyset$

Taxonomy & Taxonomic Constraints

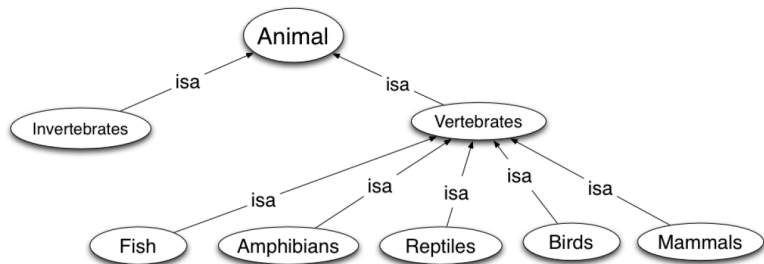
A **taxonomy** is a set of biological taxa (concepts) in an *ISA* hierarchy with *taxonomic constraints*:



- ▶ Non-Emptiness: $\text{Birds} \neq \emptyset$
- ▶ Sibling-Disjointness: $\text{Birds} \cap \text{Fish} = \emptyset$

Taxonomy & Taxonomic Constraints

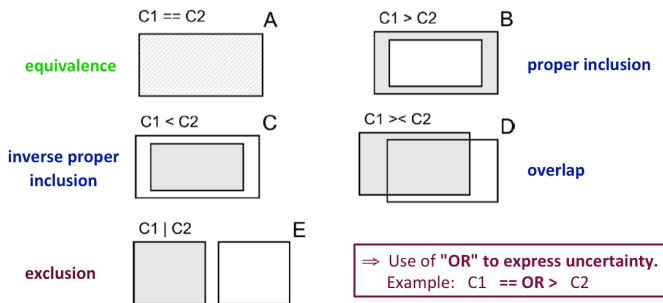
A **taxonomy** is a set of biological taxa (concepts) in an *ISA* hierarchy with *taxonomic constraints*:



- ▶ Non-Emptiness: $\text{Birds} \neq \emptyset$
- ▶ Sibling-Disjointness: $\text{Birds} \cap \text{Fish} = \emptyset$
- ▶ Coverage: $\text{Animal} = \text{Invertebrates} \cup \text{Vertebrates}$

RCC-5 Relations

Region Connection Calculus (RCC) [Ran92] abstractly describes regions (or their topological spaces) by their possible relations to each other. In EULER/X, we articulate a set of RCC-5 relations between different taxa.

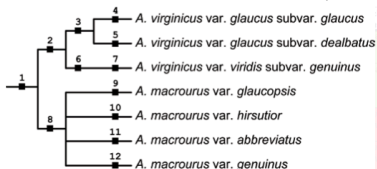


Connecting Hackel 1889 and Small 1933

Step 1: Transcribe two concept hierarchies...

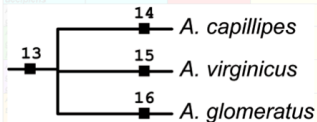
...and add unique IDs

Hackel 1889 (1-12)



| Small 1933 | Hackel 1889 |
|----------------------|--|
| <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glauca</i> subvar. <i>glauca</i> |
| <i>A. capillipes</i> | <i>A. virginicus</i> var. <i>glauca</i> subvar. <i>dealbatus</i> |
| <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>A. virginicus</i> | <i>A. virginicus</i> var. <i>viridis</i> subvar. <i>genuinus</i> |
| <i>A. glomeratus</i> | <i>A. macrourus</i> var. <i>glaucoptis</i> |
| <i>A. glomeratus</i> | <i>A. macrourus</i> var. <i>hirsutior</i> |
| <i>A. glomeratus</i> | <i>A. macrourus</i> var. <i>abbreviatus</i> |
| <i>A. glomeratus</i> | <i>A. macrourus</i> var. <i>genuinus</i> |

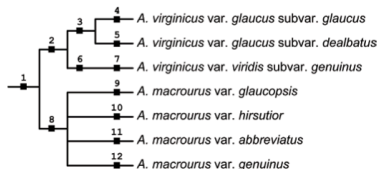
Small 1933 (13-16)



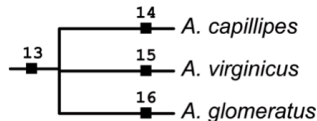
Connecting Hackel 1889 and Small 1933

Step 2: Create a table with all concept labels

Hackel 1889 (1-12)



Small 1933 (13-16)

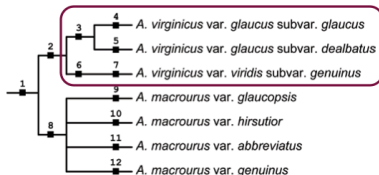


| ID | Name_Simple | According_To | Rank |
|----|--|--------------|------------|
| 1 | Andropogon | Hackel 1889 | Genus |
| 2 | Andropogon virginicus | Hackel 1889 | Species |
| 3 | Andropogon virginicus var. glaucus | Hackel 1889 | Variety |
| 4 | Andropogon virginicus var. glaucus subvar. glaucus | Hackel 1889 | Subvariety |
| 5 | Andropogon virginicus var. glaucus subvar. dealbatus | Hackel 1889 | Subvariety |
| 6 | Andropogon virginicus var. viridis | Hackel 1889 | Variety |
| 7 | Andropogon virginicus var. viridis subvar. genuinus | Hackel 1889 | Subvariety |
| 8 | Andropogon macrourus | Hackel 1889 | Species |
| 9 | Andropogon macrourus var. glaucopsis | Hackel 1889 | Variety |
| 10 | Andropogon macrourus var. hirsutior | Hackel 1889 | Variety |
| 11 | Andropogon macrourus var. abbreviatus | Hackel 1889 | Variety |
| 12 | Andropogon macrourus var. genuinus | Hackel 1889 | Variety |
| 13 | Andropogon | Small 1933 | Genus |
| 14 | Andropogon capillipes | Small 1933 | Species |
| 15 | Andropogon virginicus | Small 1933 | Species |
| 16 | Andropogon glomeratus | Small 1933 | Species |

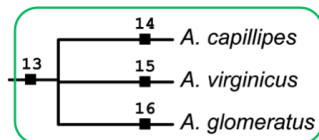
Connecting Hackel 1889 and Small 1933

Step 2: Create a table with all concept labels

Hackel 1889 (1-12)



Small 1933 (13-16)

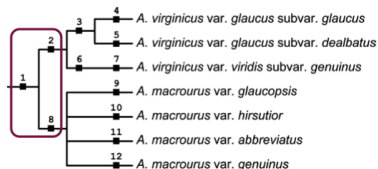


| ID | Name_Simple | According_To | Rank |
|----|--|--------------|------------|
| 1 | Andropogon | Hackel 1889 | Genus |
| 2 | Andropogon virginicus | Hackel 1889 | Species |
| 3 | Andropogon virginicus var. glaucus | Hackel 1889 | Variety |
| 4 | Andropogon virginicus var. glaucus subvar. glaucus | Hackel 1889 | Subvariety |
| 5 | Andropogon virginicus var. glaucus subvar. dealbatus | Hackel 1889 | Subvariety |
| 6 | Andropogon virginicus var. viridis | Hackel 1889 | Variety |
| 7 | Andropogon virginicus var. viridis subvar. genuinus | Hackel 1889 | Subvariety |
| 8 | Andropogon macrourus | Hackel 1889 | Species |
| 9 | Andropogon macrourus var. glaucopsis | Hackel 1889 | Variety |
| 10 | Andropogon macrourus var. hirsutior | Hackel 1889 | Variety |
| 11 | Andropogon macrourus var. abbreviatus | Hackel 1889 | Variety |
| 12 | Andropogon macrourus var. genuinus | Hackel 1889 | Variety |
| 13 | Andropogon | Small 1933 | Genus |
| 14 | Andropogon capillipes | Small 1933 | Species |
| 15 | Andropogon virginicus | Small 1933 | Species |
| 16 | Andropogon glomeratus | Small 1933 | Species |

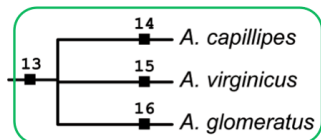
Connecting Hackel 1889 and Small 1933

Step 3: Create a table with corresponding parent/child relationships ('is_a')

Hackel 1889 (1-12)



Small 1933 (13-16)

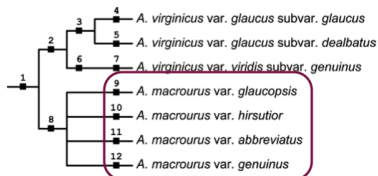


| ID | From TC | Relationship | To TC |
|----|---------|--------------|-------|
| 1 | 1 | is parent of | 2 |
| 2 | 2 | is parent of | 3 |
| 3 | 3 | is parent of | 4 |
| 4 | 3 | is parent of | 5 |
| 5 | 2 | is parent of | 6 |
| 6 | 6 | is parent of | 7 |
| 7 | 1 | is parent of | 8 |
| 8 | 8 | is parent of | 9 |
| 9 | 8 | is parent of | 10 |
| 10 | 8 | is parent of | 11 |
| 11 | 8 | is parent of | 12 |
| 12 | 13 | is parent of | 14 |
| 13 | 13 | is parent of | 15 |
| 14 | 13 | is parent of | 16 |

Connecting Hackel 1889 and Small 1933

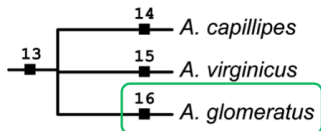
Step 4: Create a table with a suitable set of articulations

Hackel 1889 (1-12)



| ID | From_TC | Relationship_Symbol | To_TC | Rel_AccTo |
|----|------------------|---------------------|---------|--------------|
| 1 | 1 | == | 13 | Weakley 2005 |
| 2 | 2 | < | 13 | Weakley 2005 |
| 3 | 2 | > | 14 | Weakley 2005 |
| 4 | 2 | > | 15 | Weakley 2005 |
| 5 | 2 | == | 14 + 15 | Weakley 2005 |
| 6 | 3 | == | 14 | Weakley 2005 |
| 7 | 4 | < | 14 | Weakley 2005 |
| 8 | 5 | < | 14 | Weakley 2005 |
| 9 | 4 + 5 | == | 14 | Weakley 2005 |
| 10 | 6 | == | 15 | Weakley 2005 |
| 11 | 7 | == | 15 | Weakley 2005 |
| 12 | 8 | == | 16 | Weakley 2005 |
| 13 | 9 | < | 16 | Weakley 2005 |
| 14 | 10 | < | 16 | Weakley 2005 |
| 15 | 11 | < | 16 | Weakley 2005 |
| 16 | 12 | < | 16 | Weakley 2005 |
| 17 | 9 + 10 + 11 + 12 | == | 16 | Weakley 2005 |

Small 1933 (13-16)



| Small 1933 | Hackel 1889 |
|--------------------|--|
| A. capillipes | A. virginicus var. glaucus subvar. glaucus |
| A. capillipes | A. virginicus var. glaucus subvar. dealbatus |
| A. virginicus | A. virginicus var. viridis subvar. genuinus |
| A. virginicus | A. virginicus var. viridis subvar. genuinus |
| A. virginicus | A. virginicus var. viridis subvar. genuinus |
| A. glomeratus | A. macrourus var. glaucopsis |
| A. glomeratus | A. macrourus var. hirsutior |
| A. glomeratus | A. macrourus var. abbreviatus |
| A. glomeratus | A. macrourus var. genuinus |
| 2 species, 3 vars. | 7 species, 7 vars. |

Translation
Congruence

Challenges when connecting two taxonomies

- ▶ Input of articulations (Goal: achieve a complete and consistent mappings)
 - ▶ Taxonomic experts will not input ∞ articulations
 - ▶ Taxonomic experts will miss relevant articulations
 - ▶ Taxonomic experts could posit **logically inconsistent articulations**

Challenges when connecting two taxonomies

- ▶ Input of articulations (Goal: achieve a complete and consistent mappings)
 - ▶ Taxonomic experts will not input ∞ articulations
 - ▶ Taxonomic experts will miss relevant articulations
 - ▶ Taxonomic experts could posit **logically inconsistent articulations**
- ▶ Getting a merged taxonomy based on the input taxonomies and articulations

Challenges when connecting two taxonomies

- ▶ Input of articulations (Goal: achieve a complete and consistent mappings)
 - ▶ Taxonomic experts will not input ∞ articulations
 - ▶ Taxonomic experts will miss relevant articulations
 - ▶ Taxonomic experts could posit **logically inconsistent articulations**
- ▶ Getting a merged taxonomy based on the input taxonomies and articulations
- ▶ “**EULER/X Toolkit**” is being developed to explore solutions to these challenges

EULER/X Toolkit

EULER/**X**: **X** stands for its underlying reasoner

EULER/X Toolkit

EULER/**X**: **X** stands for its underlying reasoner

- ▶ EULER/FO: underlying reasoner Prover9/Mace4

EULER/X Toolkit

EULER/**X**: **X** stands for its underlying reasoner

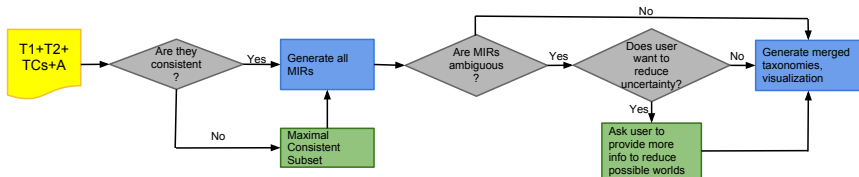
- ▶ EULER/FO: underlying reasoner Prover9/Mace4
- ▶ EULER/ASP: underlying reasoner DLV or Potassco

EULER/X Toolkit

EULER/**X**: **X** stands for its underlying reasoner

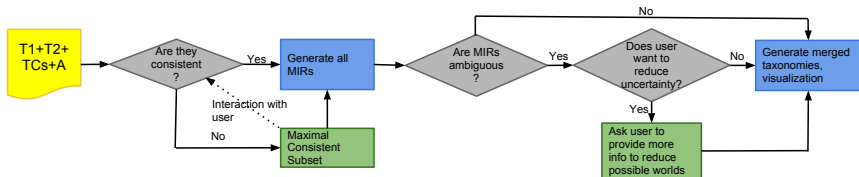
- ▶ EULER/FO: underlying reasoner Prover9/Mace4
- ▶ EULER/ASP: underlying reasoner DLV or Potassco
- ▶ EULER/PyRCC: underlying reasoner PyRCC8

EULER/X Workflow



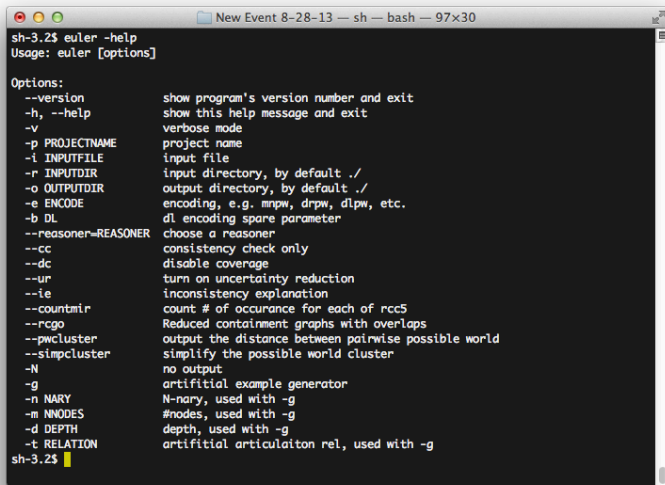
Overview of the EULER/X workflow

EULER/X Workflow



Overview of the EULER/X workflow

EULER/X Command Line Interface

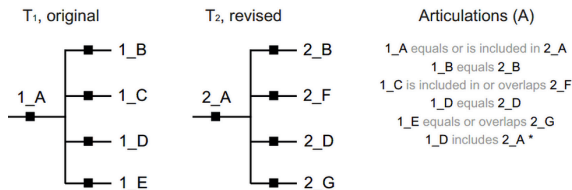


```
sh-3.25$ euler -help
Usage: euler [options]

Options:
--version          show program's version number and exit
-h, --help        show this help message and exit
-v               verbose mode
-p PROJECTNAME    project name
-i INPUTFILE      input file
-r INPUTDIR       input directory, by default ./
-o OUTPUTDIR      output directory, by default ./
-e ENCODE         encoding, e.g. mnpw, drpw, dlpw, etc.
-b DL             dl encoding spare parameter
--reasoner=REASONER choose a reasoner
--cc              consistency check only
--dc              disable coverage
--ur              turn on uncertainty reduction
--ie              inconsistency explanation
--countmir        count # of occurrence for each of rcc5
--rcgo            Reduced containment graphs with overlaps
--pwcluster       output the distance between pairwise possible world
--simplcluster    simplify the possible world cluster
-N               no output
-g               artifitial example generator
-n NARY           N-nary, used with -g
-m NNODES         #nodes, used with -g
-d DEPTH          depth, used with -g
-t RELATION       artifitial articulation rel, used with -g

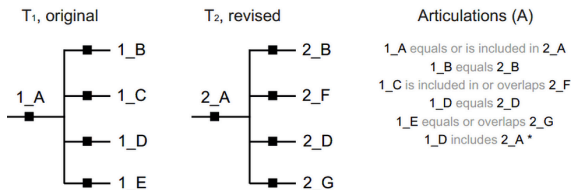
sh-3.25$
```

Synthetic Example



Two succeeding taxonomies (T_1 , T_2) and a set of expert-asserted input articulations (A) among respective taxonomic concepts.

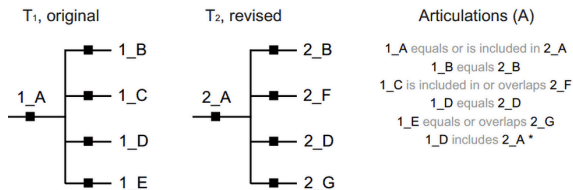
Synthetic Example



Two succeeding taxonomies (T_1 , T_2) and a set of expert-asserted input articulations (A) among respective taxonomic concepts.

Objective: Get a merged taxonomy (or several possible merged taxonomies)

EULER/X Input



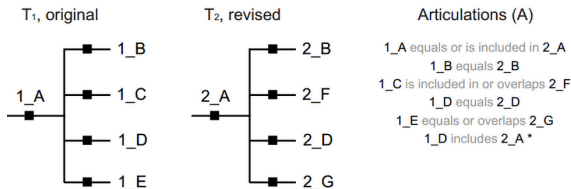
```

demo.txt
taxonomy 1 Taxonomy1
(a b c d e)

taxonomy 2 Taxonomy2
(a b d f g)

articulation synthetic example]
[1.a {is_included_in equals} 2.a]
[1.b equals 2.b]
[1.c {is_included_in overlaps} 2.f]
[1.d equals 2.d]
[1.e {equals overlaps} 2.g]
[1.d includes 2.a]
  
```

Consistency Checking / Inconsistency Repairing



Input is inconsistent!!

Remedial measure: remove [1.d includes 2.a]

Articulation <1.d includes 2.a> is inconsistent with [<1.a {is_included_in equals} 2.a>, <1.b equals 2.b>, <1.c {is_included_in overlaps} 2.f>, <1.d equals 2.d>, <1.e {equals overlaps} 2.g>]

EULER/X points out that articulation “1_D includes 2_A” is inconsistent with other articulations and should be removed.



Taxonomy Merge

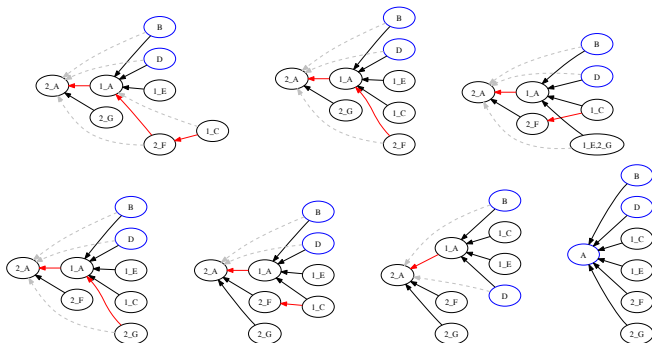
```

sh-3.25 euler -i demo.txt -e mpw
Possible world 0: {1.d^~2.d, 1.e^~2.g, 1.b^~2.b, 1.d^~2.a, 1.c^~2.a, 1.c^~2.f, 1.b^~2.a, 1.e^~2.a, 1.e^~2.a, 1.a^~2.b, 1.a^~2.d, 1.a^~2.g,
1.a^~2.f, 1.d^12.g, 1.d^12.b, 1.e^12.f, 1.e^12.d, 1.e^12.b, 1.c^12.d, 1.c^12.b, 1.c^12.g, 1.b^12.d, 1.b^12.g, 1.b^12.f}
Possible world 1: {1.d^~2.d, 1.b^~2.b, 1.d^~2.a, 1.c^~2.a, 1.b^~2.a, 1.e^~2.a, 1.a^~2.a, 1.a^~2.b, 1.a^~2.d, 1.e^~2.f, 1.e^~2.g, 1.a^~2.g,
1.c^~2.g, 1.c^~2.f, 1.d^12.f, 1.d^12.g, 1.d^12.b, 1.e^12.d, 1.e^12.b, 1.c^12.d, 1.c^12.b, 1.b^12.d, 1.b^12.g, 1.b^12.f}
Possible world 3: {1.d^~2.d, 1.b^~2.b, 1.d^~2.a, 1.c^~2.a, 1.c^~2.f, 1.b^~2.a, 1.e^~2.a, 1.a^~2.a, 1.a^~2.b, 1.a^~2.d, 1.e^~2.f, 1.e^~2.g,
1.a^~2.f, 1.d^12.f, 1.d^12.g, 1.d^12.b, 1.e^12.d, 1.e^12.b, 1.c^12.d, 1.c^12.b, 1.c^12.g, 1.b^12.d, 1.b^12.g, 1.b^12.f}
Possible world 4: {1.d^~2.d, 1.b^~2.b, 1.d^~2.a, 1.c^~2.a, 1.b^~2.a, 1.e^~2.a, 1.a^~2.a, 1.a^~2.b, 1.a^~2.d, 1.e^~2.f, 1.e^~2.g, 1.e^~2.f, 1.e^~2.g,
1.c^~2.g, 1.c^~2.f, 1.d^12.f, 1.d^12.g, 1.d^12.b, 1.e^12.d, 1.e^12.b, 1.c^12.d, 1.c^12.b, 1.b^12.d, 1.b^12.g, 1.b^12.f}
Possible world 5: {1.d^~2.d, 1.b^~2.b, 1.e^~2.a, 1.d^~2.a, 1.c^~2.a, 1.b^~2.a, 1.e^~2.a, 1.a^~2.a, 1.a^~2.b, 1.a^~2.d, 1.e^~2.f, 1.e^~2.g, 1.e^~2.f, 1.e^~2.g,
1.c^~2.g, 1.c^~2.f, 1.d^12.f, 1.d^12.g, 1.d^12.b, 1.e^12.d, 1.e^12.b, 1.c^12.d, 1.c^12.b, 1.b^12.d, 1.b^12.g, 1.b^12.f}
Possible world 6: {1.d^~2.d, 1.b^~2.b, 1.d^~2.a, 1.c^~2.a, 1.c^~2.f, 1.b^~2.a, 1.e^~2.a, 1.a^~2.a, 1.a^~2.b, 1.a^~2.d, 1.e^~2.f, 1.e^~2.g, 1.e^~2.f, 1.e^~2.g,
1.c^~2.g, 1.c^~2.f, 1.d^12.f, 1.d^12.g, 1.d^12.b, 1.e^12.d, 1.e^12.b, 1.c^12.d, 1.c^12.b, 1.c^12.g, 1.b^12.d, 1.b^12.g, 1.b^12.f}

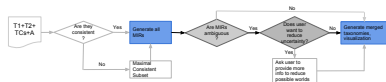
sh-3.25 more demo_mir.csv
1.o,2.a,{, <}
1.o,2.b,>
1.o,2.d,>
1.o,2.g,{, >}
1.o,2.f,{, >}
1.c,2.a,<
1.c,2.b,!
1.c,2.d,!
1.c,2.g,{, >}
1.c,2.f,{, <}
1.b,2.a,<
1.b,2.b,=
1.b,2.d,!
1.b,2.g,!
1.b,2.f,!
1.e,2.a,<
1.e,2.b,!
1.e,2.d,!
1.e,2.g,{, >}
1.e,2.f,{, >}
1.d,2.a,<
1.d,2.b,!
1.d,2.d,=
1.d,2.g,!
1.d,2.f,!
sh-3.25
    
```



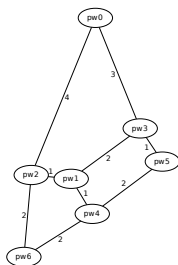
Taxonomy Merge



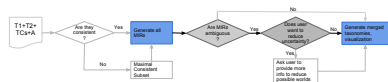
Set of possible worlds (pw0, pw1, . . . , pw6)
resulting from the repaired input example



Visual clustering of similar possible worlds



Visualization of possible worlds, where the distance between two possible worlds is the shortest distance traceable in the graph (e.g., the distance between worlds 5 and 6 is 4).



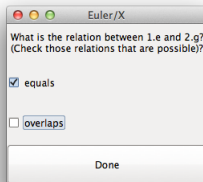
Uncertainty Reduction

```
sh-3.2$ euler -i demo.txt -e mnpw --ur
```



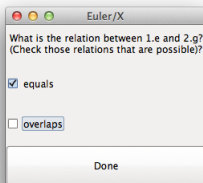
Uncertainty Reduction

```
sh-3.2$ euler -i demo.txt -e mmpw --ur
```

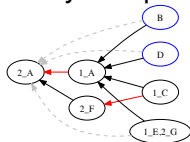


Uncertainty Reduction

```
sh-3.2$ euler -i demo.txt -e mnpw --ur
```



will give you only one possible world:

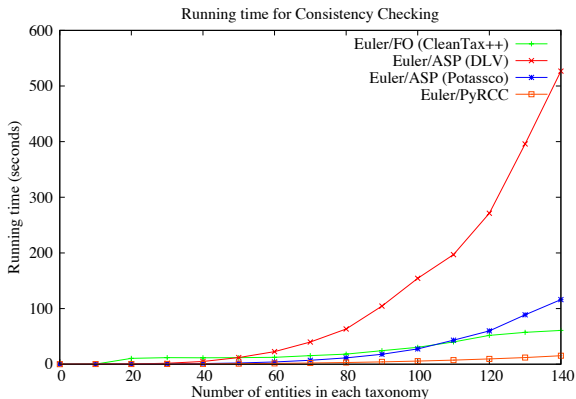


Demo – from end to end

<http://www.youtube.com/watch?v=i3JlbZ5ehQY&feature=youtu.be>

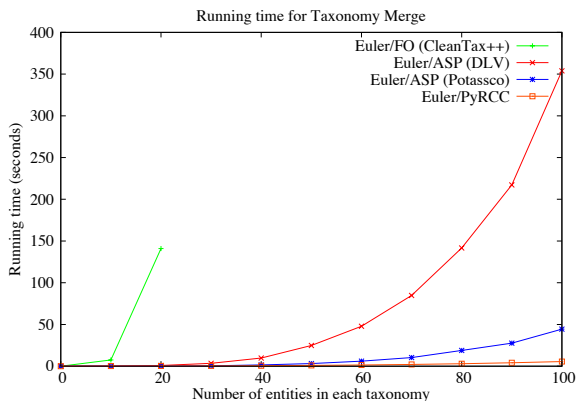
17/24

Performance Results – Consistency Checking



Running time for consistency checking on synthetic taxonomies

Performance Results – Merging Taxonomies



Running time for merging taxonomies on synthetic taxonomies.

Performance Results – Summary

- ▶ How do different reasoner-based tools work?
 - ▶ EULER/FO : checks consistency by calling Mace4 once and then generates each mir by calling Prover9 (for $m * n$ mir's assuming there are m, n entities in each taxonomy)
 - ▶ others: only invoke the reasoner once to check consistency and merge taxonomies

Performance Results – Summary

- ▶ How do different reasoner-based tools work?
 - ▶ EULER/FO : checks consistency by calling Mace4 once and then generates each mir by calling Prover9 (for $m * n$ mir's assuming there are m, n entities in each taxonomy)
 - ▶ others: only invoke the reasoner once to check consistency and merge taxonomies
- ▶ Consistency Checking
 - ▶ EULER/FO is slower than EULER/ASP (Potassco) when the number of entities in each taxonomy is less than 100, but faster when it is more than 100.

Performance Results – Summary

- ▶ How do different reasoner-based tools work?
 - ▶ EULER/FO : checks consistency by calling Mace4 once and then generates each mir by calling Prover9 (for $m * n$ mir's assuming there are m, n entities in each taxonomy)
 - ▶ others: only invoke the reasoner once to check consistency and merge taxonomies
- ▶ Consistency Checking
 - ▶ EULER/FO is slower than EULER/ASP (Potassco) when the number of entities in each taxonomy is less than 100, but faster when it is more than 100.
- ▶ Merging Taxonomy
 - ▶ EULER/PyRCC > EULER/ASP (Potassco) >> EULER/ASP (DLV) >> EULER/FO (Prover9/Mace4)
 - ▶ Note: coverage constraints cannot be asserted using PyRCC8.

Comparison with Ontology Matching

| | Ontology Matching | Taxonomy Integration |
|--------|---|---|
| Input | ontologies (OWL, RDF), similarity / distance vector | taxonomies, articulations & taxonomic constraints |
| Output | finding the matching between entities; getting merged ontologies. | finding the relation between entities; getting merged taxonomies. |

Related Work & Contribution

Related Work

- ▶ MoReTax
 - first attempt of using RCC-5 in taxonomy articulations
- ▶ CLEAN TAX by Dave Thau.
 - prior work of EULER/FO

Related Work & Contribution

Related Work

- ▶ MoReTax
 - first attempt of using RCC-5 in taxonomy articulations
- ▶ CLEAN TAX by Dave Thau.
 - prior work of EULER/FO

Contribution

- ▶ Encodings
- ▶ Provenance
- ▶ Visualization
- ▶ The first real-life implementation of ASP in a domain

Encodings – Some ideas

1. Enumerate all the minterms (e.g. $ABC\bar{D}$...):

%%% minterms

$r(M) :- \#int(M), M \geq 1, M \leq \#maxint.$

$in(X, M) \vee out(X, M) :- r(M), concept(X).$

$vr(M) \vee ir(M) :- r(M).$

2. Encode the constraints (including articulations) on the minterms:

%%% $c1_b$ isa $c1_a$

$ir(X) :- in(c1_b, X), out(c1_a, X).$

$:- \#countX: vrs(X), in(c1_b, X), in(c1_a, X) = 0.$

3. Decode according to the validity of the minterms:

$hint(X, Y, 0) :- concept(X), concept(Y), vr(R), in(X, R), out(Y, R).$

$hint(X, Y, 1) :- concept(X), concept(Y), vr(R), in(X, R), in(Y, R).$

$hint(X, Y, 2) :- concept(X), concept(Y), vr(R), out(X, R), in(Y, R).$

$rel(X, Y, "=") :- not hint(X, Y, 0), hint(X, Y, 1), not hint(X, Y, 2).$

...

Q&A

Thanks!