



SemMatcher: A Tool for Matching Ontology-based Schemas

Carlos Eduardo Pires (cesp@cin.ufpe.br)

Damires Souza (damires@ifpb.edu.br)

Thiago Pachêco (tpap@cin.ufpe.br)

Ana Carolina Salgado (acs@cin.ufpe.br)

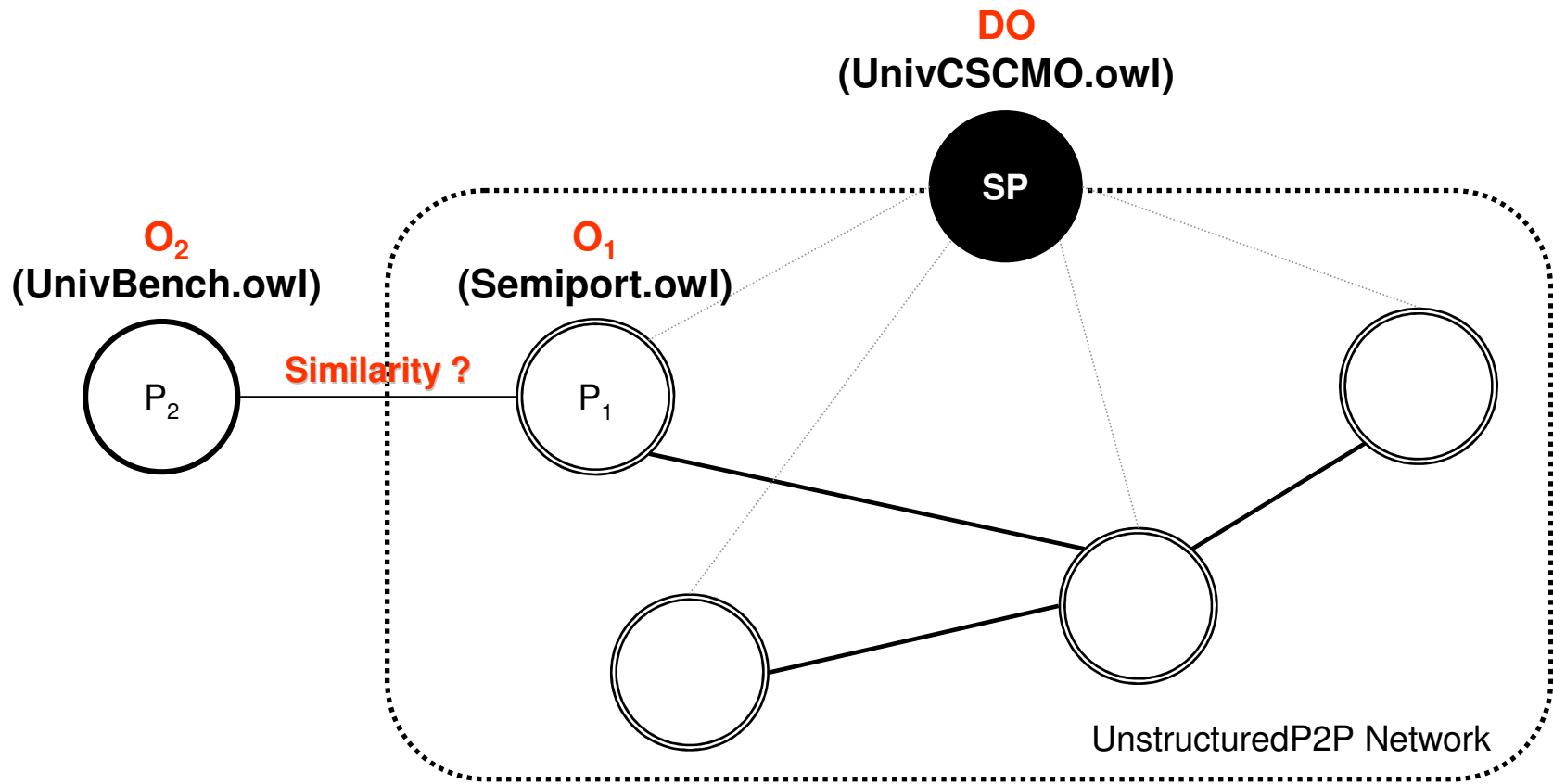


Outline

- ▶ Working Scenario
- ▶ Ontology Matching Process
 - ▶ Semantic matching
 - ▶ Combination, ranking and selection
 - ▶ Global similarity measure
 - ▶ Final alignment
- ▶ SemMatcher
- ▶ Conclusions



Working Scenario



Semantic Community of Peers

** Education Domain **

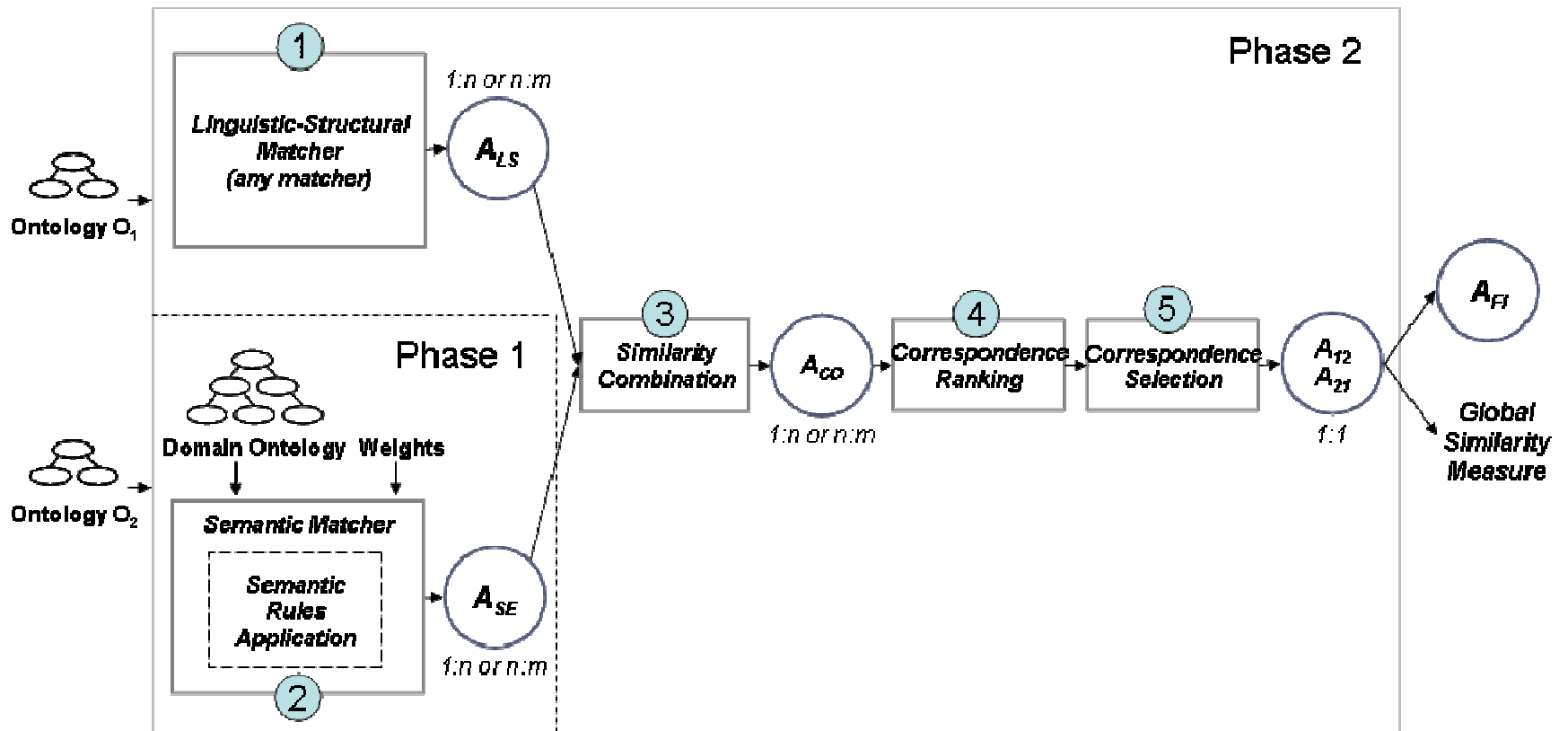


Goal

- ▶ Propose a semantic-based ontology matching tool which has been instantiated in a PDMS
- ▶ Contributions
 - ▶ Identification of semantic correspondences between two peer ontologies
 - ▶ Taking into account a domain ontology as background knowledge
 - ▶ Determination of the global similarity between two peer ontologies

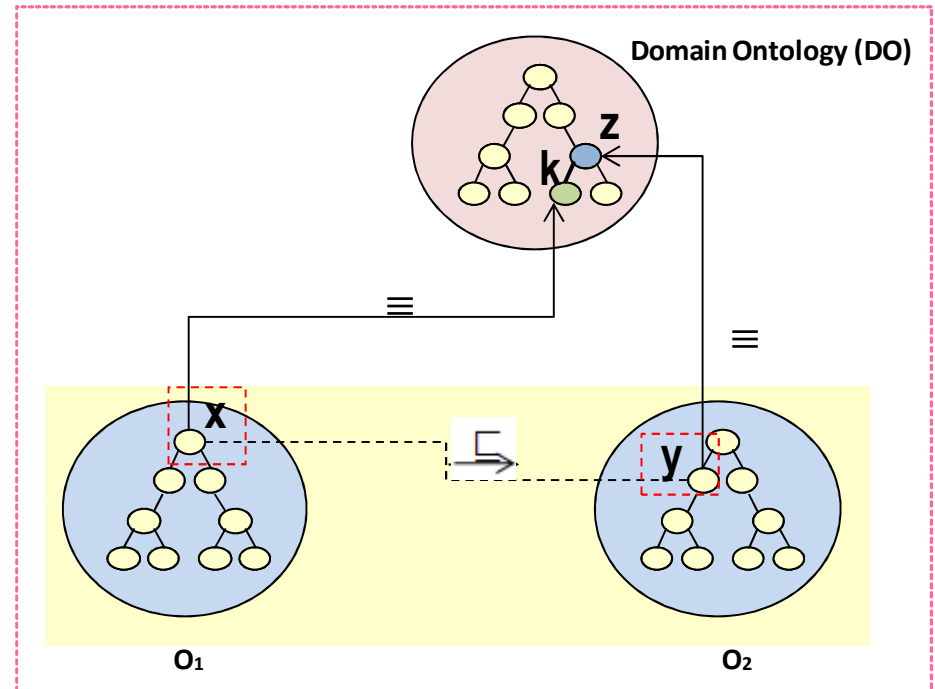


Ontology Matching Process



Semantic matching

- ▶ **Domain Ontology (DO)**
 - ▶ Reliable reference available on the Web
 - ▶ Used as Background Knowledge
 - ▶ Bridge the conceptual differences or similarities between two peer ontologies



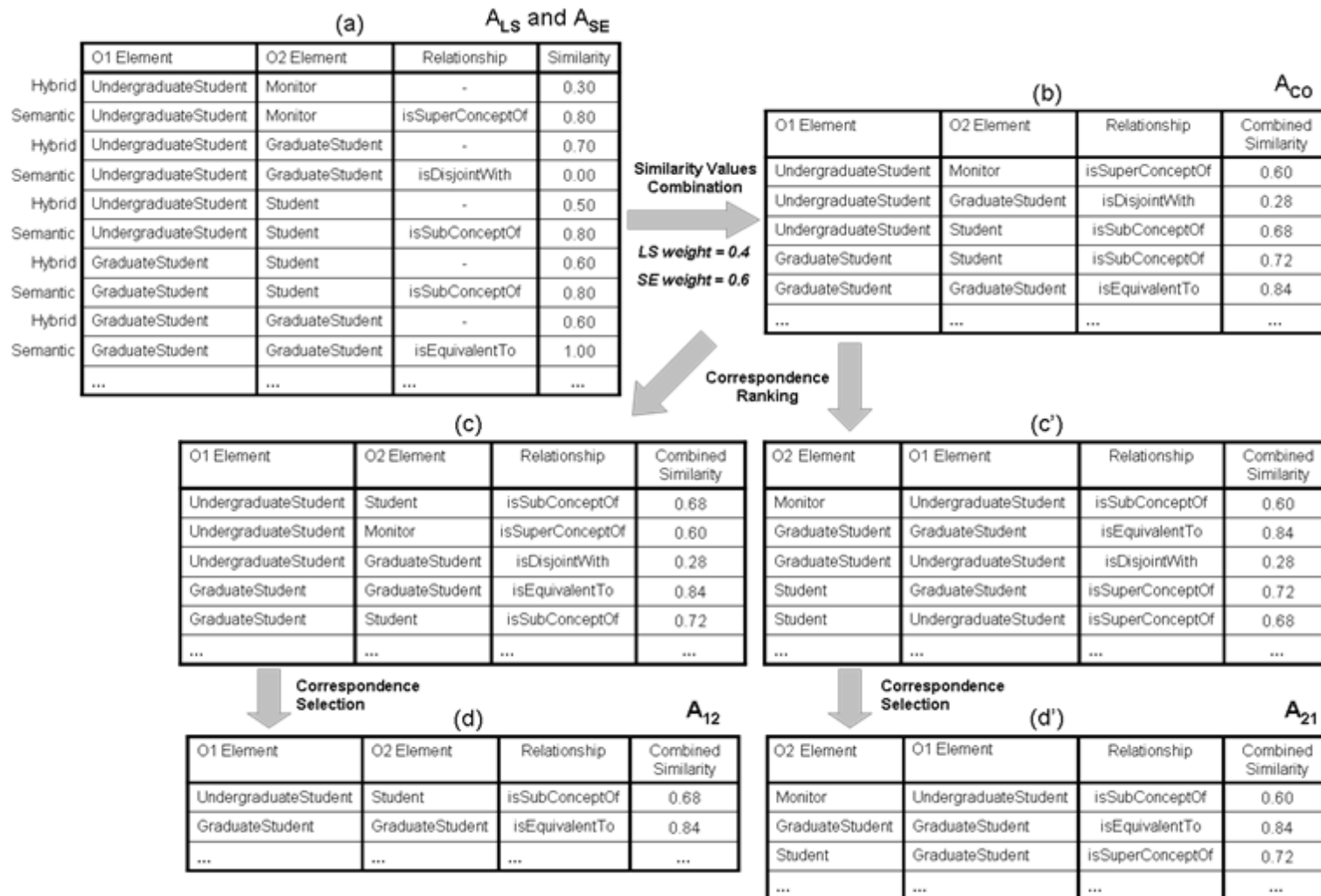
Semantic matching

- ▶ **Definition.** A semantic correspondence is represented by one of the following expressions:
 - ▶ $O_1:x \equiv O_2:y$, an *isEquivalentTo* correspondence
 - ▶ $O_1:x \sqsubseteq O_2:y$, an *isSubConceptOf* correspondence
 - ▶ $O_1:x \sqsupseteq O_2:y$, an *isSuperConceptOf* correspondence
 - ▶ $O_1:x \triangleright O_2:y$, an *isPartOf* correspondence
 - ▶ $O_1:x \triangleleft O_2:y$, an *isWholeOf* correspondence
 - ▶ $O_1:x \approx O_2:y$, an *isCloseTo* correspondence
 - ▶ $O_1:x \perp O_2:y$, an *isDisjointWith* correspondence

(*) x and y are elements belonging to the peer ontologies



Combination, ranking and selection



Global similarity measure

- ▶ Calculated using the intermediate outputs: A_{12} and A_{21}
- ▶ The result is a value corresponding to the similarity between 2 peers
- ▶ The size of an ontology ($|O|$) is determined by the

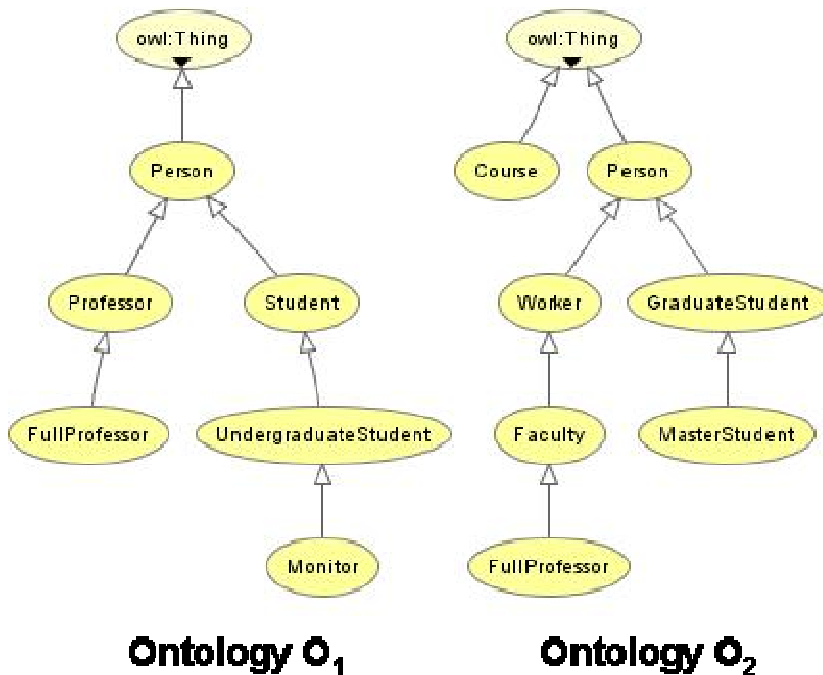
$$Dice(O_1, O_2) = \frac{|A_{12}| + |A_{21}|}{|O_1| + |O_2|} \quad Weighted(O_1, O_2) = \frac{\sum_{i=1}^{|A_{12}|} n + \sum_{i=1}^{|A_{21}|} n}{|O_1| + |O_2|} \quad Overlap(O_1, O_2) = \frac{|A_{12} \cap A_{21}|}{\min(|O_1|, |O_2|)}$$

[Aümuller et al., 2005] [Castano et al., 1998] [Rijsbergen, 1979]



Global similarity measure

$$\text{Weighted Average}(O_1, O_2) = \frac{\sum_{i=1}^{|A_{12}|} n + \sum_{j=1}^{|A_{21}|} n}{|O_1| + |O_2|}$$



Alignment A₁₂

- (1, Person, Person, isEquivalentTo, 1.0)
- (2, FullProfessor, FullProfessor, isEquivalentTo, 1.0)
- (3, UndergraduateStudent, Course, isPartOf, 0.3)
- (4, Student, Person, isSubConceptOf, 0.8)
- (5, Professor, Faculty, isSubConceptOf, 0.8)

Alignment A₂₁

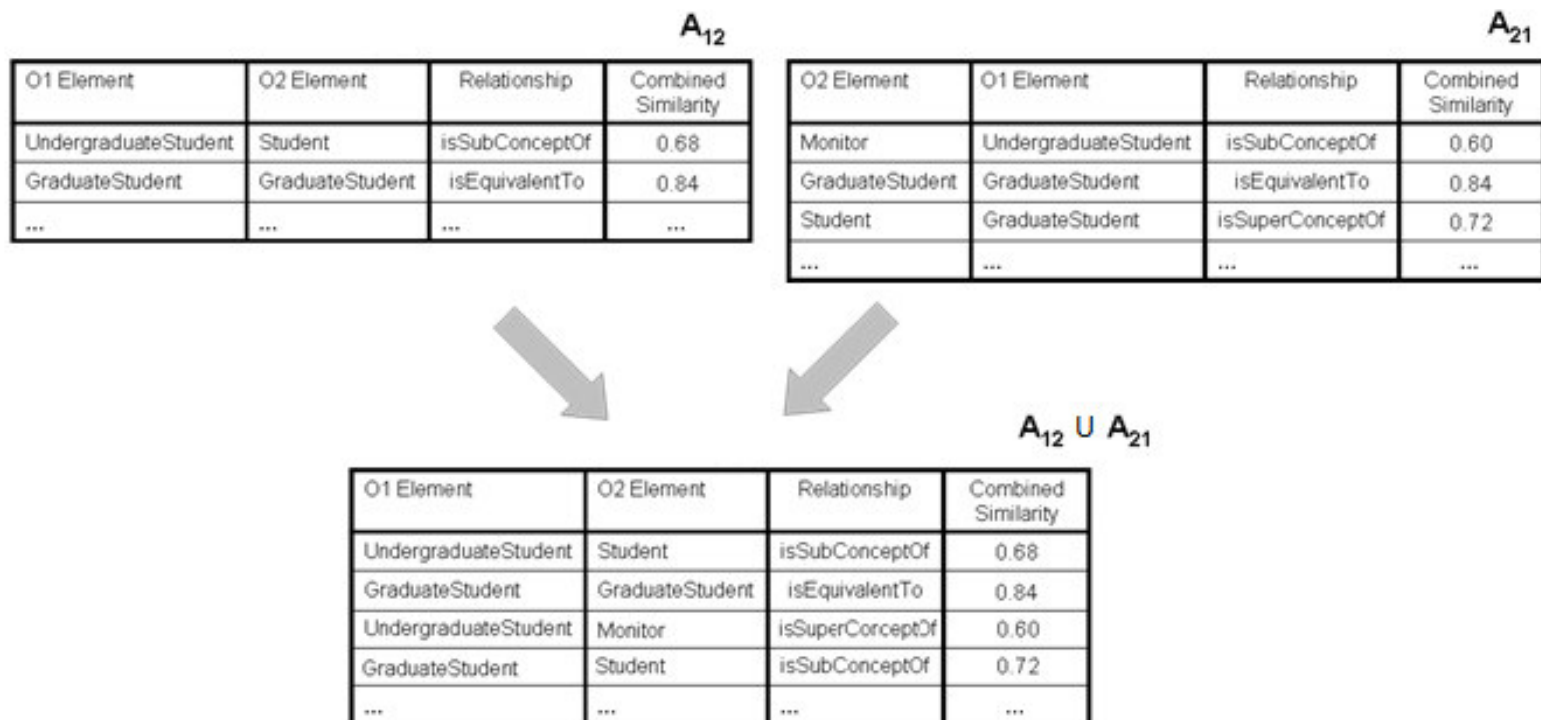
- (1, Person, Person, isEquivalentTo, 1.0)
- (2, FullProfessor, FullProfessor, isEquivalentTo, 1.0)
- (3, Course, UndergraduateStudent, isWholeOf, 0.3)
- (4, Worker, Person, isSubConceptOf, 0.8)
- (5, GraduateStudent, UndergraduateStudent, isDisjointWith, 0.0)
- (6, Faculty, Professor, isSuperConceptOf, 0.8)
- (7, MasterStudent, Student, isSubConceptOf, 0.8)

$$\text{Weighted Average}(O_1, O_2) = \frac{(1.0 + 1.0 + 0.3 + 0.8 + 0.8) + (1.0 + 1.0 + 0.3 + 0.8 + 0.0 + 0.8 + 0.8)}{|6| + |7|} = 0.66$$



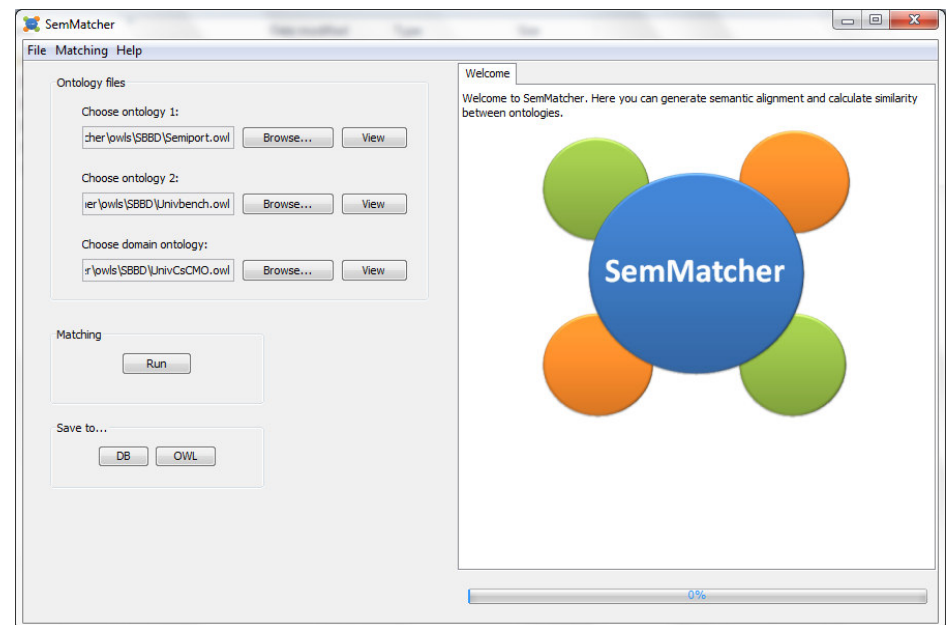
Final alignment

- ▶ Union between A_{12} and A_{21}
- ▶ Result:
 - ▶ Semantically enriched alignment



SemMatcher

- ▶ Implemented using:
 - ▶ Java
 - ▶ Jena
 - ▶ OWL API
- ▶ Allows to execute:
 - ▶ H-Match
 - ▶ Alignment API
- ▶ Supports our algorithm

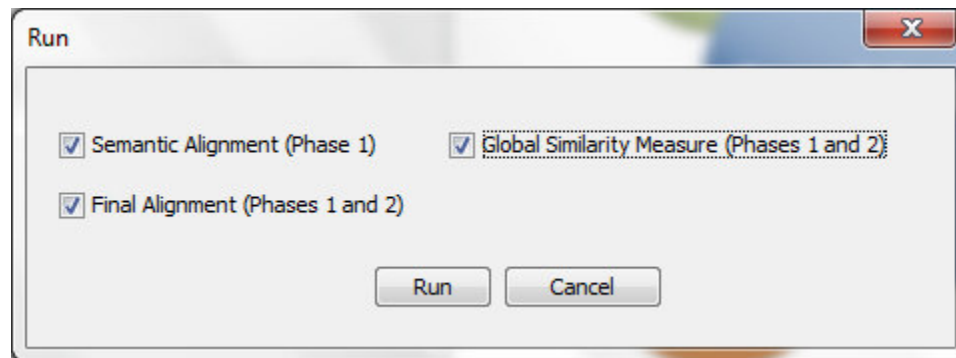


SemMatcher's main screen



SemMatcher

- ▶ Choose the phases to execute



SemMatcher

► Semantic alignment

The screenshot shows the SemMatcher application window. The interface is divided into several sections:

- Ontology files:** Three sections for selecting ontologies. The first section is for 'Choose ontology 1:' with the file path 'e:\owls\SBBD\Semiport.owl'. The second is for 'Choose ontology 2:' with the file path 'e:\owls\SBBD\Univbench.owl'. The third is for 'Choose domain ontology:' with the file path 'e:\owls\SBBD\UnivCsCMO.owl'. Each section has 'Browse...' and 'View' buttons.
- Matching:** A 'Run' button.
- Save to...:** 'DB' and 'OWL' buttons.
- Table:** A table showing semantic alignments between two ontologies. The table has four columns: 'Ontology 1', 'Correspondence', 'Ontology 2', and 'Measure'. The table is currently displaying 25 rows of results.

Ontology 1	Correspondence	Ontology 2	Measure
AdministrativeStaff	isEquivalentTo	AdministrativeStaff	1
AdministrativeStaff	isSubConceptOf	Worker	0,8
AdministrativeStaff	isSuperConceptOf	Director	0,8
AdministrativeStaff	isSuperConceptOf	Dean	0,8
AdministrativeStaff	isSuperConceptOf	Chair	0,8
AdministrativeStaff	isCloseTo	Faculty	0,7
AdministrativeStaff	isCloseTo	Assistant	0,7
Article	isEquivalentTo	Article	1
Article	isSubConceptOf	Publication	0,8
Article	isSuperConceptOf	JournalArticle	0,8
Article	isSuperConceptOf	ConferencePaper	0,8
Article	isSuperConceptOf	TechnicalReport	0,8
Article	isPartOf	Book	0,3
Article	isCloseTo	Manual	0,7
Article	isCloseTo	UnofficialPublication	0,7
Article	isCloseTo	Specification	0,7
Article	isCloseTo	Book	0,7
Article	isCloseTo	Software	0,7
AssistantProfessor	isSubConceptOf	Professor	0,8
AssistantProfessor	isPartOf	Course	0,3
AssistantProfessor	isPartOf	ResearchProject	0,3
AssistantProfessor	isCloseTo	VisitingProfessor	0,7
AssistantProfessor	isDisjointWith	AssociateProfessor	0
AssistantProfessor	isDisjointWith	FullProfessor	0
Book	isEquivalentTo	Book	1
Book	isSubConceptOf	Publication	0,8

100% - Done

SemMatcher

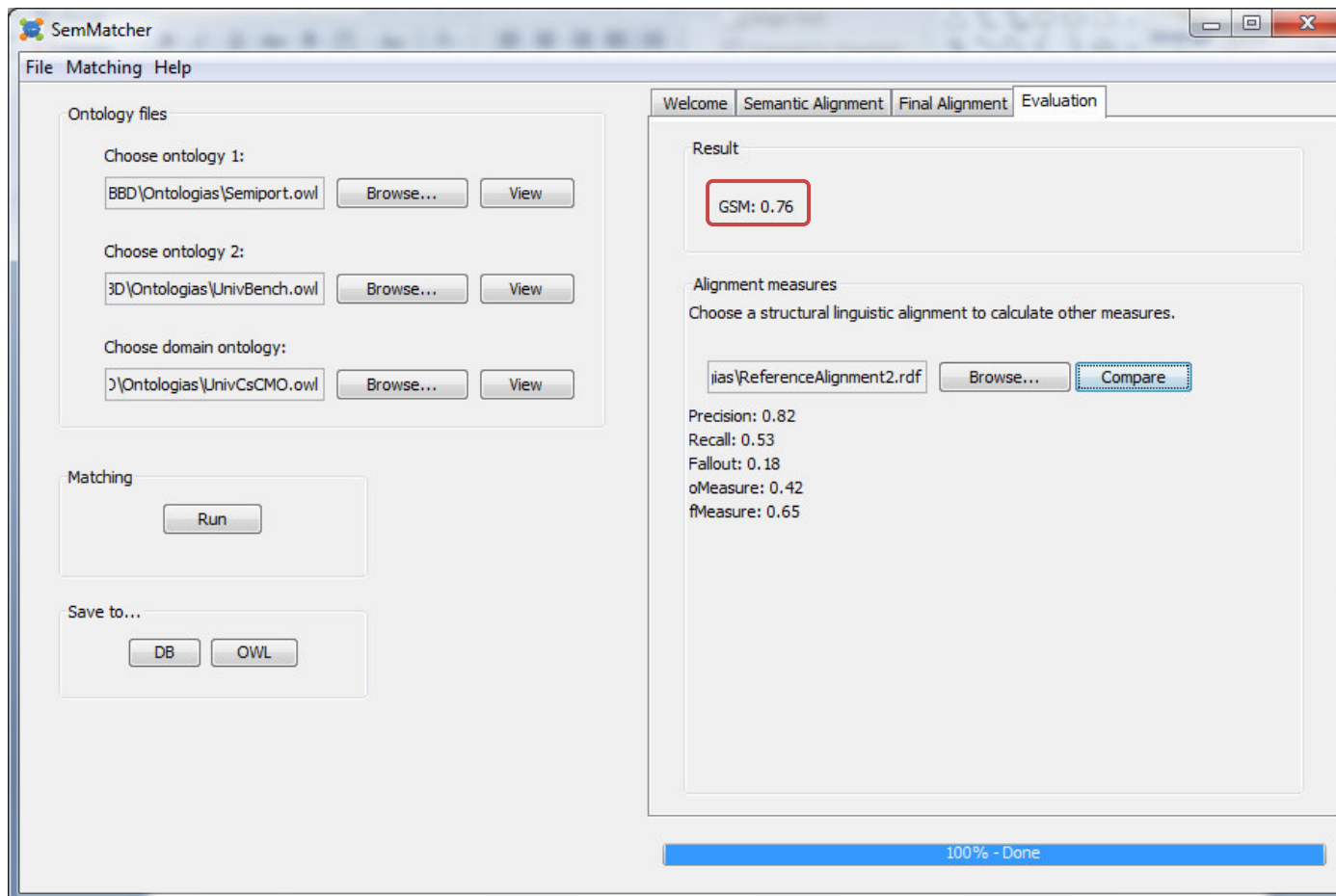
- ▶ Final alignment = $A_{12} \cup A_{21}$

The screenshot shows the SemMatcher application window. The 'Final Alignment' tab is active, displaying a table of correspondences between two ontologies. The table has four columns: 'Ontology 1', 'Correspondence', 'Ontology 2', and 'Measure'. The table lists various classes and their relationships, such as 'AdministrativeStaff' being equivalent to 'AdministrativeStaff' with a measure of 1, or 'AdministrativeStaff' being a super-concept of 'Director' with a measure of 0,68. The status bar at the bottom indicates '100% - Done'.

Ontology 1	Correspondence	Ontology 2	Measure
AdministrativeStaff	isEquivalentTo	AdministrativeStaff	1
AdministrativeStaff	isSuperConceptOf	Director	0,68
AdministrativeStaff	isSuperConceptOf	Dean	0,68
AdministrativeStaff	isSuperConceptOf	Chair	0,68
Article	isEquivalentTo	Article	1
Article	isSuperConceptOf	JournalArticle	0,84
Article	isSuperConceptOf	ConferencePaper	0,68
AssistantProfessor	isSubConceptOf	Professor	0,84
AssistantProfessor	No semantic relation	ResearchAssistant	0,28
AssistantProfessor	No semantic relation	TeachingAssistant	0,28
Book	isEquivalentTo	Book	1
ClericalStaff	isSubConceptOf	AdministrativeStaff	0,81
Conference	No semantic relation	ConferencePaper	0,3
Department	isEquivalentTo	Department	1
DevelopmentProject	isCloseTo	ResearchProject	0,75
DoctoralThesis	No semantic relation	MasterStudent	0,23
Event	No semantic relation	Department	0,23
Faculty	isEquivalentTo	Faculty	1
Faculty	isSuperConceptOf	PostDoc	0,68
FullProfessor	isEquivalentTo	FullProfessor	1
FullProfessor	isSubConceptOf	Professor	0,84
FullProfessor	isCloseTo	VisitingProfessor	0,75
GraduateStudent	isEquivalentTo	GraduateStudent	1
GraduateStudent	isSuperConceptOf	MasterStudent	0,81
GraduateStudent	isPartOf	Course	0,42
GraduateStudent	No semantic relation	GraduateCourse	0,25

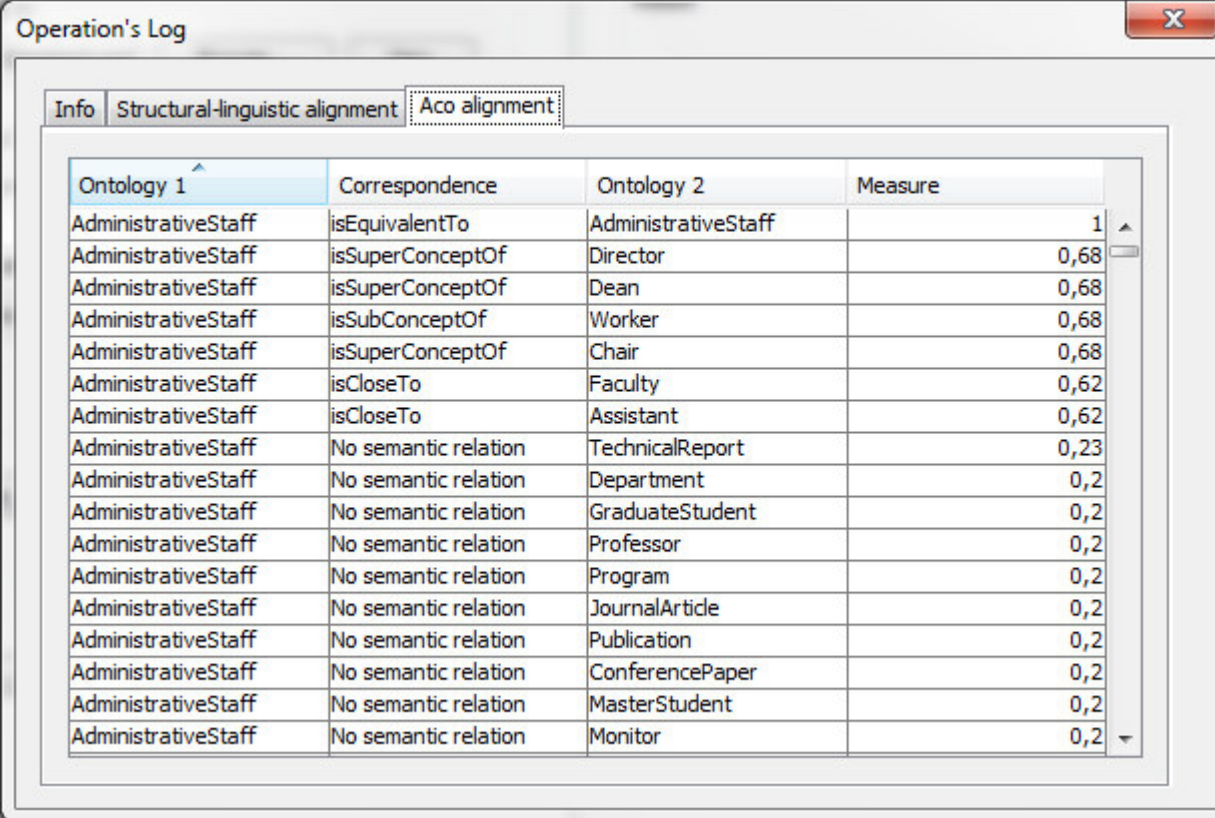
SemMatcher

▶ Global similarity measure



SemMatcher

- ▶ Log: Intermediate steps



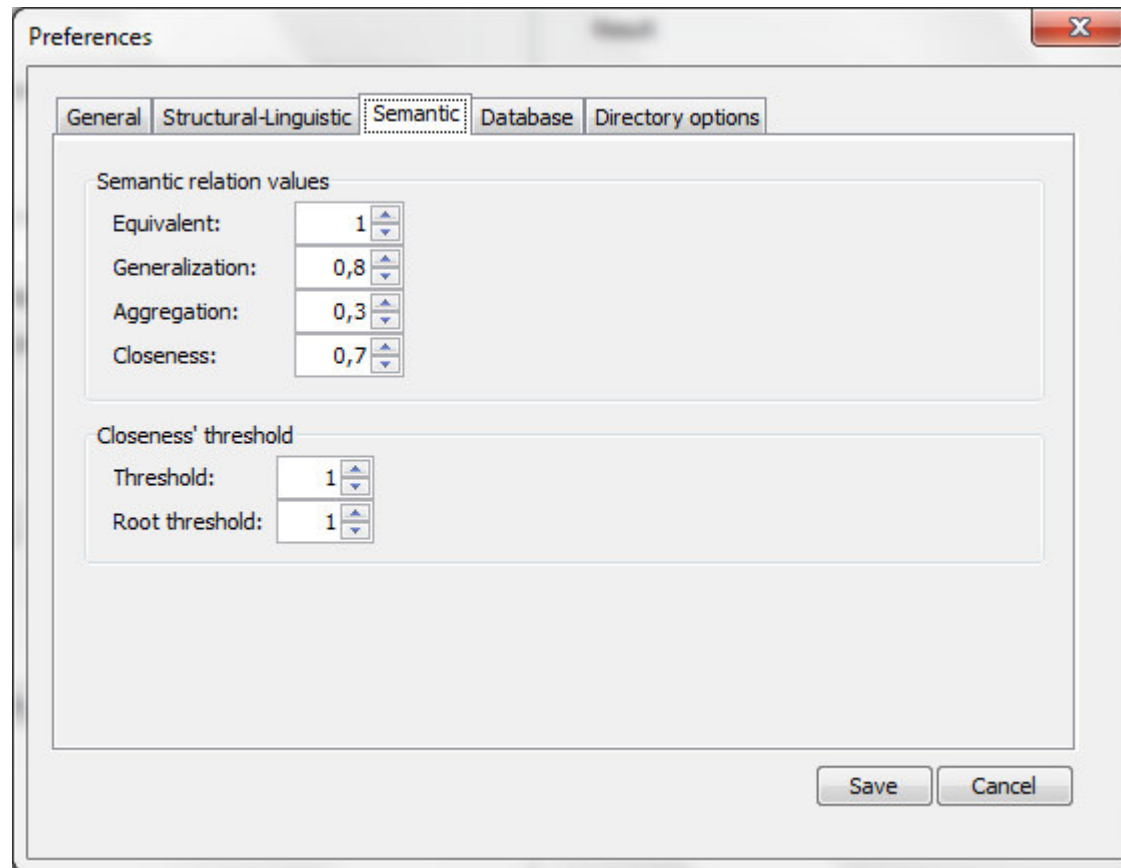
The screenshot shows a window titled "Operation's Log" with a close button (X) in the top right corner. The window contains three tabs: "Info", "Structural-linguistic alignment", and "Aco alignment". The "Aco alignment" tab is active and displays a table with the following columns: "Ontology 1", "Correspondence", "Ontology 2", and "Measure". The table lists 18 rows of alignments, starting with "AdministrativeStaff" in Ontology 1 and various terms in Ontology 2, with corresponding measure values ranging from 0,2 to 1.

Ontology 1	Correspondence	Ontology 2	Measure
AdministrativeStaff	isEquivalentTo	AdministrativeStaff	1
AdministrativeStaff	isSuperConceptOf	Director	0,68
AdministrativeStaff	isSuperConceptOf	Dean	0,68
AdministrativeStaff	isSubConceptOf	Worker	0,68
AdministrativeStaff	isSuperConceptOf	Chair	0,68
AdministrativeStaff	isCloseTo	Faculty	0,62
AdministrativeStaff	isCloseTo	Assistant	0,62
AdministrativeStaff	No semantic relation	TechnicalReport	0,23
AdministrativeStaff	No semantic relation	Department	0,2
AdministrativeStaff	No semantic relation	GraduateStudent	0,2
AdministrativeStaff	No semantic relation	Professor	0,2
AdministrativeStaff	No semantic relation	Program	0,2
AdministrativeStaff	No semantic relation	JournalArticle	0,2
AdministrativeStaff	No semantic relation	Publication	0,2
AdministrativeStaff	No semantic relation	ConferencePaper	0,2
AdministrativeStaff	No semantic relation	MasterStudent	0,2
AdministrativeStaff	No semantic relation	Monitor	0,2



SemMatcher

► Configuration



Conclusions

- ▶ Our matching tool tries to overcome limitations of linguistic and structural approaches by using **domain ontologies** as **background knowledge**
 - ▶ A **semantic matcher** identifies, besides traditional types of correspondences, other ones (e.g., closeness and disjointness)
 - ▶ Determination of a **global similarity measure** between two ontologies (not only between their elements)
- ▶ Further work
 - ▶ Include an alignment-reuse **matcher**





SemMatcher: A Tool for Matching Ontology-based Schemas

Carlos Eduardo Pires (cesp@cin.ufpe.br)

Damires Souza (damires@ifpb.edu.br)

Thiago Pachêco (tpap@cin.ufpe.br)

Ana Carolina Salgado (acs@cin.ufpe.br)

