Consensus Ontology Generation in a Socially Interacting MultiAgent System

Ergun Biçici

Koç University, Istanbul, Turkey
www.ku.edu.tr

September 20, 2006

ESAS 2006 Presentation

Summary

- An approach for building consensus ontologies from ontologies of a network of socially interacting agents
- Each agent has its own conceptualization of the world.
- Interactions between agents based on queries and their assessments allow us to model the quality of resources.
- The dynamic emergence of consensus mimics the evolution of language.
- An algorithm for generating the consensus ontologies using the authoritative agent's conceptualization is presented.
- Consensus ontologies are evaluated by using heuristic measures of similarity based on the component ontologies.

Today's Talk

- Reaching Consensus and Introduction
- Challenges and Contributions
- Formal Definitions
  - Problem Formulation
  - Mapping Concepts
- Social Networks
- Algorithm, Experiments, and Results
- Conclusion

Reaching Consensus

- An *ontology* is a thesaurus [5], which answers the question of "what there is" [4] in a domain.
- Semantic agreement among heterogeneous ontologies is not always possible.
- Various ontologies for the same domain exists. Which one is the best?
- We address the problem of building consensus ontologies from multiple heterogeneous ontologies belonging to a number of agents interacting with each other.

Sample ontology from the data set.

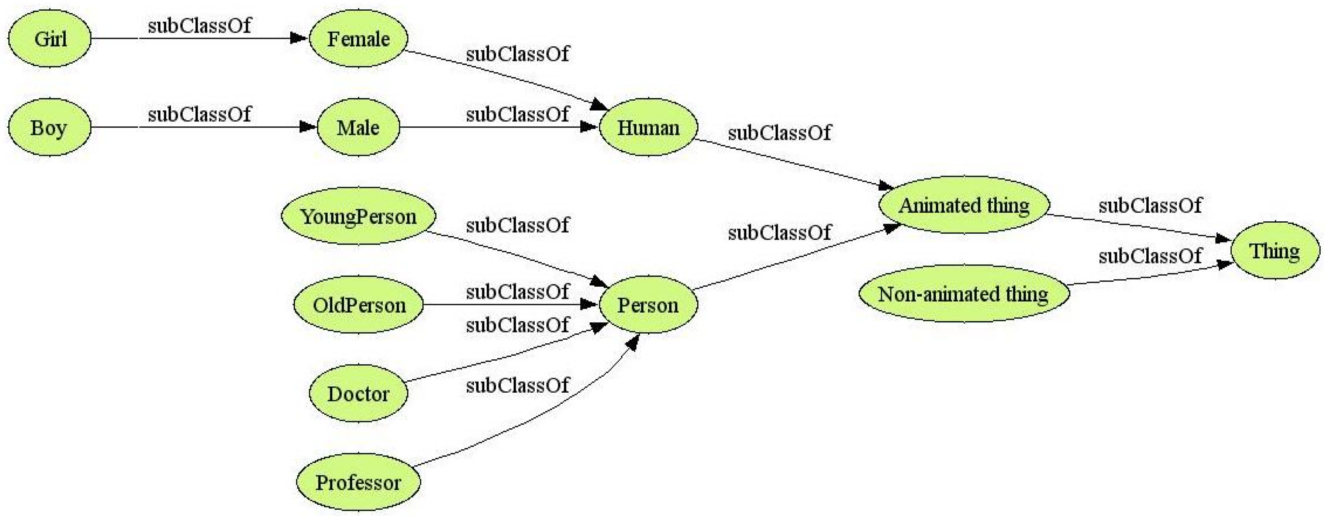


Figure 1: Example ontology for the domain “humans”.

Motivation and Challenges

- Forming a consensus ontology is important because it provides:
  - A vocabulary to which agents can refer when misunderstandings occur.
  - A unified world view supported by the members, which facilitates distributed knowledge management.

- Technical Challenges:
  - *Conceptual mapping:* We need to be able to find mappings between different ontologies.
    $\ggg$ We present a method of concept mapping based on the conceptual structures in the ontologies.

– *Consensus generation:* What is a good way to generate the consensus ontology? How to resolve conflicting statements?
$\ggg$ Build the consensus ontology based on combining the beliefs of experts in each domain where expertise is gained by agents through social interactions.

– *Consensus evaluation:* How can we measure the goodness of the final consensus?
$\ggg$ We developed heuristic measures for evaluating the consensus ontology based on three different perspectives.

Related Work

- Majority based reinforcement formulation done by Stephens and Huhns [6].
- Query transformation and semantic agreement detection done by Aberer *et. al.* [2].
- Ontological mediator (oracle) based terminological mismatch resolution done by Campbell and Shapiro [1].
- Noy [3] discussed techniques for finding correspondences between ontologies.
- Building consensus ontology applied in service composition [7].

Problem Formulation

- An ontology is a 2-tuple $< \mathcal{C}, <_{\mathcal{C}} >$ where $\mathcal{C}$ is the set of concepts and $<_{\mathcal{C}}$ is the "subClassOf" relation.

- $C_1 <_{\mathcal{C}} C_2 \; \Leftrightarrow \; C_1$ is a subconcept of C_2.

- A multiagent system (MAS) is a set of agents, $\mathcal{A} = \{A_1 \ldots A_n\}$.

- Each agent A_i has an ontology $\mathcal{O}_i = < \mathcal{C}_i, <_{\mathcal{C}_i} >$ and a lexicon, $\mathcal{L}_i$, which defines the set of allowable terms.

- $O^{\cap}$ denotes the intersection ontology and $<_{\cap}$ defines a consistent ordering with all of the given set of orderings.

Heuristic Measures

- Lexical:

(1) $SM(L_i, L_j) := max\left(0, \frac{min(|L_i|,|L_j|) - ed(L_i,L_j)}{min(|L_i|,|L_j|)}\right)$

(2) $\overline{SM}(\mathcal{L}_1, \mathcal{L}_2) := \frac{1}{|\mathcal{L}_1|}\sum_{L_i \in \mathcal{L}_1} max_{L_j \in \mathcal{L}_2} SM(L_i, L_j)$

(3) $\overline{\overline{SM}}(\mathcal{L}_C, \mathbf{L}) = \frac{1}{|\mathbf{L}|}\sum_{\mathcal{L}_i \in \mathbf{L}} \frac{\overline{SM}(\mathcal{L}_C,\mathcal{L}_i) + \overline{SM}(\mathcal{L}_i,\mathcal{L}_C)}{2}$

- Conceptual:

(4) $AS(C_i, <_\mathcal{C}) := \{C_j \in \mathcal{C} \mid C_i \ <_\mathcal{C} \ C_j \ \vee \ C_i = C_j\}$

(5) $TS(L, \mathcal{O}_i, \mathcal{O}_j) := \begin{cases} TS_1(L, \mathcal{O}_i, \mathcal{O}_j), & \text{if } L \in \mathcal{L}_j \\ TS_2(L, \mathcal{O}_i, \mathcal{O}_j), & \text{if } L \notin \mathcal{L}_j \end{cases}$

(6) $\overline{TS}(\mathcal{O}_i, \mathcal{O}_j) := \frac{1}{|\mathcal{L}_i|}\sum_{L \in \mathcal{L}_i} TS(L, \mathcal{O}_i, \mathcal{O}_j)$

(7) $\overline{\overline{TS}}(\mathcal{O}, \mathbf{O}) = \frac{1}{|\mathbf{O}|}\sum_{\mathcal{O}_i \in \mathbf{O}} \frac{\overline{TS}(\mathcal{O},\mathcal{O}_i) + \overline{TS}(\mathcal{O}_i,\mathcal{O})}{2}$

Heuristic Measures

- Information retrieval:

(8) $precision(\mathcal{O}, \mathcal{O}_C) = \frac{|elements(\mathcal{O}) \cap elements(\mathcal{O}_C)|}{elements(\mathcal{O})}$

(9) $recall(\mathcal{O}, \mathcal{O}_C) = \frac{|elements(\mathcal{O}) \cap elements(\mathcal{O}_C)|}{elements(\mathcal{O}_C)}$

(10) $FMeasure(\mathcal{O}, \mathcal{O}_C) = \frac{2 \times recall(\mathcal{O}) \times precision(\mathcal{O})}{recall(\mathcal{O}) + precision(\mathcal{O})}$

(11) $\overline{Precision}(\mathcal{O}_C) = \frac{1}{|\mathbf{O}|} \sum_{\mathcal{O}_i \in \mathbf{O}} precision(\mathcal{O}_i)$

(12) $\overline{Recall}(\mathcal{O}_C) = \frac{1}{|\mathbf{O}|} \sum_{\mathcal{O}_i \in \mathbf{O}} recall(\mathcal{O}_i)$

(13) $\overline{FMeasure}(\mathcal{O}_C) = \frac{1}{|\mathbf{O}|} \sum_{\mathcal{O}_i \in \mathbf{O}} FMeasure(\mathcal{O}_i)$

(14) $O^{\cap} = < \bigcap_{i=1}^{n} \mathcal{C}_i, <_{\cap} >$

Methods for Mapping Concepts

$\mathcal{F} \subseteq \mathcal{L}_{\mathcal{C}} \times \mathcal{C}, L \in \mathcal{L}_{\mathcal{C}}, \mathcal{F}(L) = \{C \in \mathcal{C} \mid (L, C) \in \mathcal{F}\}$,
$\mathcal{F}^{-1}(C) = \{L \in \mathcal{L}_{\mathcal{C}} \mid (L, C) \in \mathcal{F}\}$.

(15) $UC(C_i, <_{\mathcal{C}}) := \{C_j \in \mathcal{C} \mid C_i <_{\mathcal{C}} C_j \vee C_i = C_j\}$

(16) $LCM(L_1, \mathcal{O}_1, L_2, \mathcal{O}_2) := CM(\mathcal{F}(L_1), \mathcal{O}_1, \mathcal{F}(L_2), \mathcal{O}_2)$

(17) $CM(C_1, \mathcal{O}_1, C_2, \mathcal{O}_2) := \frac{|\mathcal{F}_1^{-1}(UC(C_1, <_{\mathcal{C}_1})) \cap \mathcal{F}_2^{-1}(UC(C_2, <_{\mathcal{C}_2}))|}{|\mathcal{F}_1^{-1}(UC(C_1, <_{\mathcal{C}_1})) \cup \mathcal{F}_2^{-1}(UC(C_2, <_{\mathcal{C}_2}))|}$

(18) $OUC(C_i, <_{\mathcal{C}}) := \{C_j \in \mathcal{C} \mid C_i <_{\mathcal{C}} C_j \vee C_i = C_j\}_{\leqslant_{<_{\mathcal{C}}}}$

(19) $OM(A_{\leqslant_A}, B_{\leqslant_B}) = \sum_{i=1}^{n-1} a_i \leqslant_A a_{i+1} \Leftrightarrow \mathbf{m}(a_i) \leqslant_B \mathbf{m}(a_{i+1})$

(20) $OCM(C_1, \mathcal{O}_1, C_2,\ \mathcal{O}_2) := \frac{OM(OUC(C_1, <_{\mathcal{C}_1}),\ OUC(C_2, <_{\mathcal{C}_2}))}{min(|OUC(C_1, <_{\mathcal{C}_1})|,\ |OUC(C_2, <_{\mathcal{C}_2})|)}$

Concept Mapping Algorithm

The threshold levels for concept mapping are set as 0.6, 0.3, and 0.5 for α_1, α_2, and α_3 correspondingly.

Given: *Two lexical entries* L_1 *and* L_2 *belonging to ontologies* $\mathcal{O}_1$ *and* $\mathcal{O}_2$ *correspondingly, find out if their concepts do match with the thresholds* α_1, α_2, *and* α_3.

if $SM(L_1, L_2) \geqslant \alpha_1$ **then**
  if $OCM(L_1, \mathcal{O}_1, L_2, \mathcal{O}_2) \geqslant \alpha_2$ **then**
    $\mathbf{m}(L_1) = L_2$
else if $OCM(L_1, \mathcal{O}_1, L_2, \mathcal{O}_2) \geqslant \alpha_3$ **then**
  $\mathbf{m}(L_1) = L_2$
else
  $\mathbf{m}(L_1) \neq L_2$

Social Networks

- A social network is a set of agents which socially interact with each other by using queries and answers.
- Models store information about the expertise of an agent, the projected ability to produce correct answers, and their sociability, the projected ability to produce correct referrals.
- Agents pose queries based on their interests and evaluate others based on the answers they receive.
- Queries, answers, and interests are sets of $\langle term,\ expertiseValue \rangle$ tuples.

Similarity. The similarity of Q to E is found as follows:

$$Q \diamond E = \frac{\sum_i q_i \times e_j}{\sqrt{n \sum_{i=1}^{n} q_i^2}},$$

where n is the number of terms in the query, $q_i \in Q$, $e_j \in E$, and $\mathbf{m}(q_i) = e_j$.

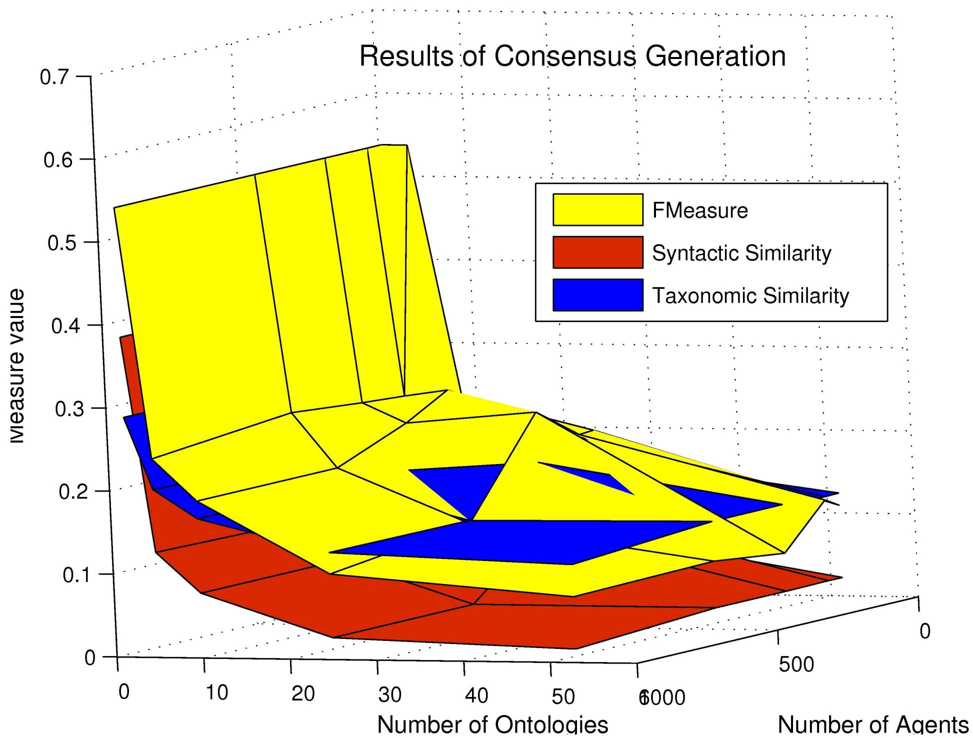
Building Consensus Based on Domain Expertise

Given: *A set of agents, $\mathcal{A}$, sharing a set of ontologies, $\mathbf{O}$, find the consensus ontology, $\mathcal{O}_C$, represented by a consistent set of statements such that it represents a consensus for the MAS.*

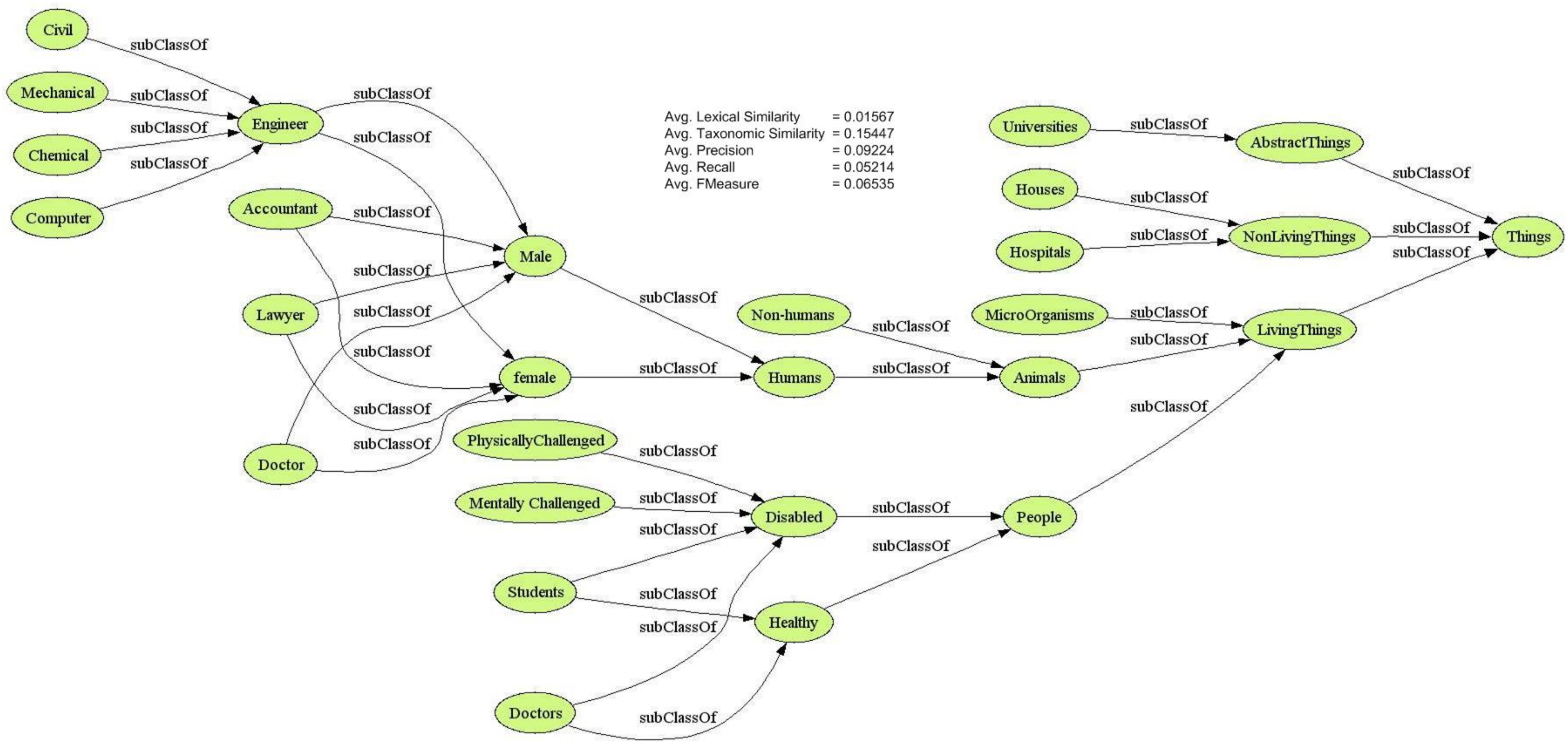
```
O_C = ⋂_{i=1}^{n} O_{A_i}
while newLeafSetSize ≠ LeafSetSize do
    LeafSet = getLeaves(O_C)
    LeafSetSize = |LeafSet|
    for C_subj ∈ LeafSet do
        A_expert = getDomainExpert(O, C_subj)
        expansionSet = getDomainConceptualization(O_{A_expert}, C_subj)
        for C_obj ∈ expansionSet do
            C'_obj = getBestMatchingConcept(O, C_obj)
            if C'_obj ≠ ∅ then
                add(O_C, C_subj, C'_obj)
            else
                add(O_C, C_subj, C_obj)
        end
        newLeafSet = getLeaves(O_C)
        newLeafSetSize = |newLeafSet|
    end
end
```

Experiments and Results

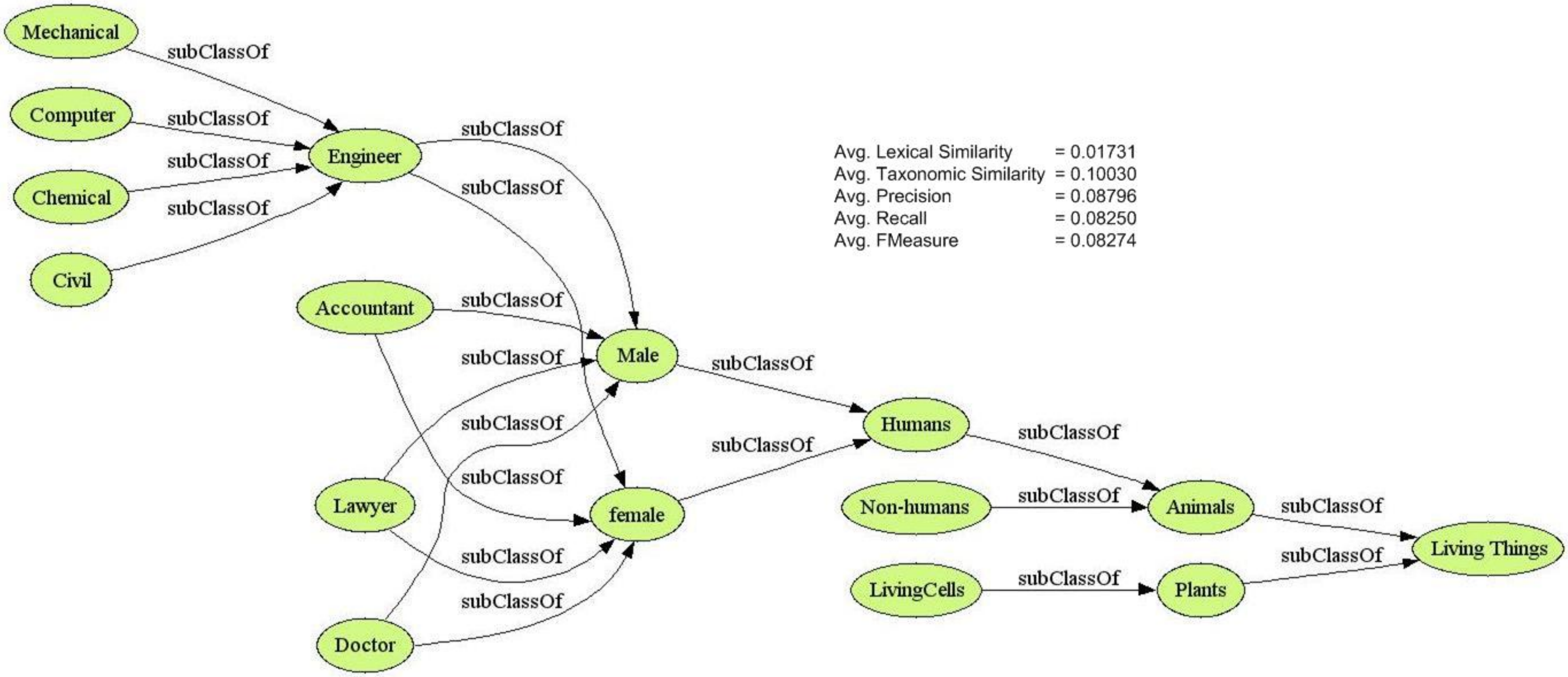
- Agents ranging from 5 to 1000, having differing ontologies ranging from 2 to 53.
- The expertise levels are initialized to the depth of the domain.



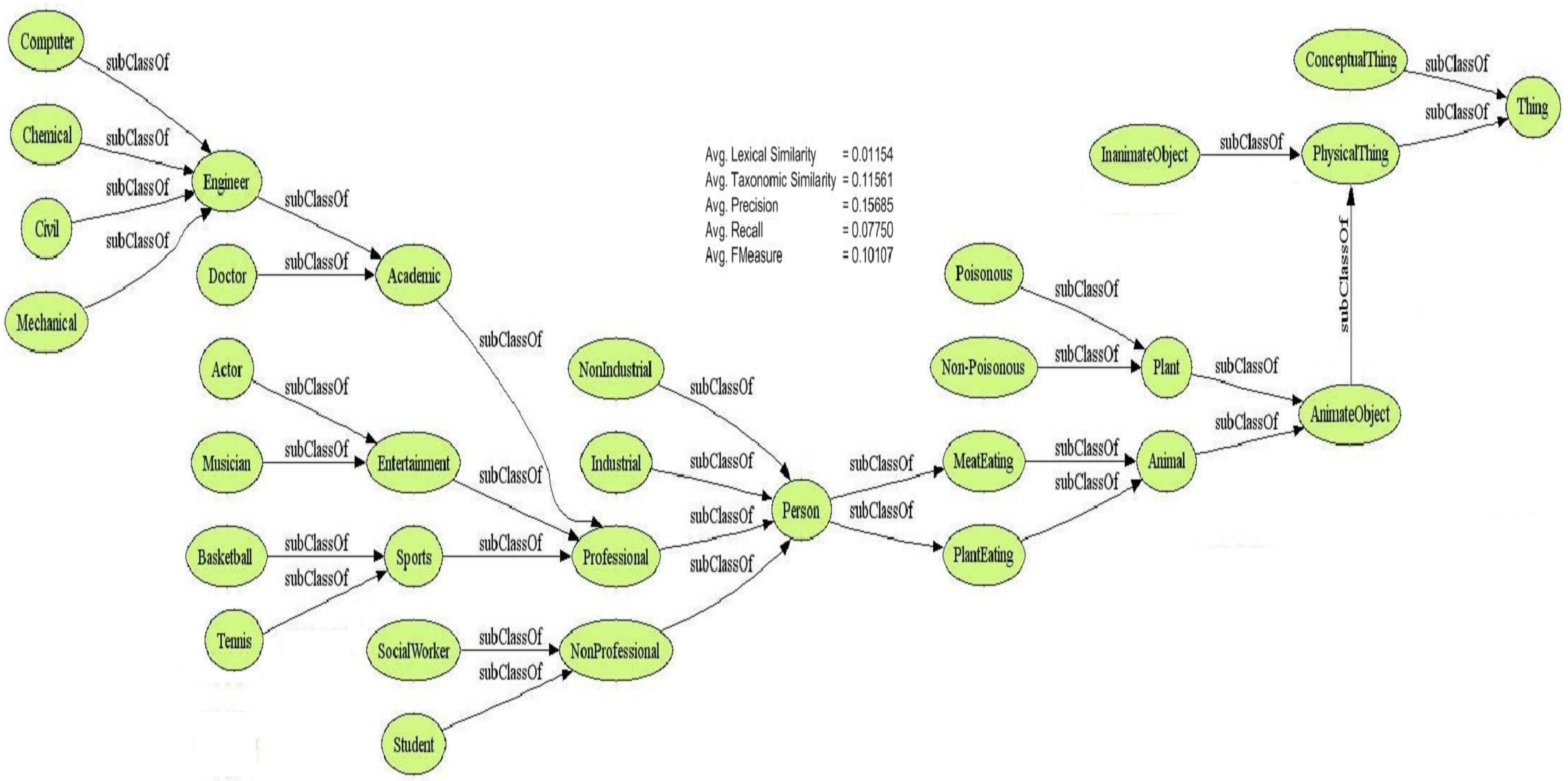
Evolution of Consensus



Evolution of Consensus (2)



Evolution of Consensus (3)



Conclusion

- We studied the generation of consensus ontologies among agents having differing ontologies in a multiagent system.
- We developed heuristics measures for evaluating the consensus ontology and methods for conceptual processing.
- Interactions between agents based on queries and their assessments allow us to model the quality of resources.
- The dynamic emergence of consensus mimics the evolution of language.
- An algorithm for generating the consensus ontologies using the authoritative agent's conceptualization is presented.

- We presented a method of concept mapping based on the conceptual structures in the ontologies.
- We expect that this research will help us understand and formalize the tradeoffs between approaches to building consensus which can later determine inference mechanisms that can be in place.

Future Work

- What happens if individual agents' ontologies are also allowed to evolve?

- Refinement of the final consensus ontology based on some heuristics (i.e. *coherence continuum*: Smoothing the consensus ontology when expert agent chosen for hierarchically ordered domains alternate so that we choose to retain the alternating agent's recommendation.)

- When choosing good domain experts, we can also check the expertise of experts in the upper levels of the domain with some added decaying function effect. Assumes that an expert agent chosen for a given concept term is likely to be good in its subconcepts.

References

[1] Alistair Campell and Stuart C. Shapiro. Algorithms for ontological mediation. In Sanda Harabagiu, editor, *Use of WordNet in Natural Language Processing Systems: Proceedings of the Conference*, pages 102–107. Association for Computational Linguistics, Somerset, New Jersey, 1998.

[2] M. Hauswirth K. Aberer, P. Cudr-Mauroux. Start making sense: The chatty web approach for global semantic agreements. *Journal of Web Semantics*, 1(1), 2003.

[3] Natalya Fridman Noy. Semantic integration: A survey of

ontology-based approaches. *SIGMOD Record*, 33(4):65–70, 2004.

[4] W. V. O. Quine. *Philosophy of Logic*. Harvard University Press, second edition, 1986.

[5] Dana S. Scott. Capturing concepts with data structures. In *Proc.DS-2, IFIP TC-2 Conference on Knowledge and Data, Portugal*, November 1986.

[6] Larry M. Stephens and Michael N. Huhns. Consensus ontologies: Reconciling the semantics of web pages and agents. *IEEE Internet Computing*, 5(5):92–95, 2001.

[7] Andrew B. Williams, Anand Padmanabhan, and M. Brian

Blake. Local consensus ontologies for b2b-oriented service composition. In *AAMAS*, pages 647–654, 2003.

Thank you!

Acknowledgments: The research reported here was supported in part by North Carolina State University. The author would like to thank Munindar P. Singh, James Lester, and Jon Doyle for helpful discussions and for their guidance and support during the term of this research.