A Strategy for Automated Meaning Negotiation in Distributed Information Retrieval

Vadim Ermolayev[1], Natalya Keberle[1], Wolf-Ekkehard Matzke[2], and Vladimir Vladimirov[1]

[1] Dept of IT, Zaporozhye National Univ., Ukraine
`eva,kenga,vvlad@zsu.zp.ua`
[2] Cadence Design Systems, GmbH, Feldkirchen, Germany
`wolf@cadence.com`

Abstract. The paper reports on the development of the formal framework to design strategies for multi-issue non-symmetric meaning negotiations among software agents in a distributed information retrieval system. The advancements of the framework are the following. A resulting strategy compares the contexts of two background domain theories not concept by concept, but the whole context to the other context by accounting the relationships among concepts, the properties, the constraints over properties, and the available instances. It contains the mechanisms for measuring contextual similarity through assessing propositional substitutions and to provide argumentation through generating extra contexts. It uses presuppositions for choosing the best similarity hypotheses and to make the mutual concession to the common sense monotonic. It provides the means to evaluate the possible eagerness to concede through semantic commitments and related notions of knowledgeability and degree of reputation.

1 Introduction

Information systems in Distributed Information Retrieval are characterized by the fact that Information Resource Providers (IRPs) and Users who pose queries form an Open System in the sense that:

- **Semantic heterogeneity**: The IRPs are legally and physically autonomous and do not care about establishing the common set of concepts describing their resources. A User's background theory of a domain may as well have various semantic discrepancies with the ontologies describing involved information resources.
- **Resource changes without notice**: IRPs normally do not notify other IRP-s and users about the changes in their resources because they may not even know about these IRP-s and Users. The changes may occur at any time and affect both resource semantics and the terms of use.

An information retrieval system in the mentioned settings should be intelligent in the following aspects. Firstly, it needs to accumulate the common knowledge describing the involved IRPs with their resources and its users with their individual terminological preferences in a uniform and coherent way. Secondly, it should pro-actively cope with the dynamic changes or reconfigurations of its collection of

registered resources affecting both semantics and the conditions of resource usage. And, thirdly, it should pro-actively conduct distributed information retrieval by keeping an eye on how different participants of the team obey their commitments to the collaboration convention while performing their parts of the query posed by a user. Normally it is achieved through striking a sort of a contract deal and monitoring contract execution.

In the humans' world mentioned aspects of collaborative intelligent behavior are often arranged through various sorts of negotiation. Hence, if a framework for automating negotiations among autonomous software components in an open information retrieval system is elaborated we'll make a substantial advancement towards solving the abovementioned problems inferred by semantic heterogeneity and dynamic changes of a resource. We'll denote negotiating intelligent software components in an information retrieval system as actors. Actors are naturally implemented by software agents and play different roles. These roles in a system with a centralized or de-centralized mediator are at least:

– a **User** (an agent assisting a human user to formulate his or her queries and to process the results of these queries) – hereafter referred to as Q
– a **Mediator** (an agent or a multi-agent system which, in compliance with [1], provides services to Users through arranging the performance of their queries by available and matching IRPs) – hereafter referred to as M
– An **IRP** (which is often an agent wrapping the Information Resource (IR) and processing queries to this IR if contracted by M) – hereafter referred to as P

It should be noted that Q, M, and Ps take part in two different types of negotiations in this scenario. Negotiations are focused on two aspects: on concept meaning and on the (commercial or contract) terms of the information delivery like the price, the terms of use, the deadline. A variety of papers provide extensive results on negotiation mechanisms dealing with the commercial aspects, e.g. [2], [3], [4]. However, negotiations on the meaning of concepts are not covered well enough in the literature. In this paper we intend to make this white spot narrower.

The reminder of the paper is structured as follows. Section 2 outlines the conceptual framework for negotiations we use in our research. Section 3 describes the example we use to illustrate our framework throughout the paper and introduces the formal notation for knowledge representation. Section 4 contains the high-level description of the proposed strategy for automated multi-issue meaning negotiation. Sections 5, 6, and 7 provide more detailed elaboration of the aspects of argumentation through propositional substitutions and contexts, making presuppositions, measuring semantic similarity of concepts respectively. Section 8 denotes semantic commitment and discusses the problem of concession in negotiation. Section 9 briefly outlines the related work. Section 10 summarizes the results and outlines our future work.

2   Interactions and Negotiations

"… perhaps the most fundamental and powerful mechanism for managing inter-agent dependencies at run-time is *negotiation* — the process by which a group of agents comes to a mutually acceptable agreement on some matter." (c.f. [3]).

Negotiation type and, therefore, the corresponding negotiation mechanism is determined by the goal, the protocol, the negotiation set, and the strategies of the participants [4].

A protocol defines the rules of encounter commonly accepted by all negotiating parties. Different protocols are applicable to different negotiation encounters which may be symmetric or non-symmetric and have different number of parties: one-to-one encounters, one-to-many encounters, and many-to-many encounters. It is evident that one-to-one protocols are easier to implement and require less communication overhead. That is why it is rational (though not always possible) to substitute more complex encounters by the sequences of one-to-one negotiations.

Symmetric negotiation settings occur when the parties of the encounter do not differ neither formally, for example by their role or in the terms of Game Theory by their part of the payoff matrix, nor informally – by their capability to influence the process, for example by their reputation. In our domain symmetric negotiations may occur among the peer-agents in a decentralized mediator (e.g., [5]). Non-symmetric negotiations are more common in practice because their settings reflect real life more adequately. Indeed, even in information retrieval a) an agent within a centralized mediator (e.g., [6]) has different roles in its negotiations with IR providing agents, b) the roles of different parties are also different – a user, a mediator, an IR provider, c) the abilities of different parties to influence the process of negotiation differ as well – a mediator agent should be more conservative in its concessions on the meaning of concepts (Section 8). Of course, the mechanisms for non-symmetric encounters, though more realistic, are more difficult to implement and subsume symmetric ones as singular cases.

A negotiation set stands for the matter or the matters on which the parties try to agree. Negotiations may be single issue and multi issue according to the number of these matters. In the real world settings different matters in multi-issue negotiations are rarely independent. Hence, a multi-issue negotiation mechanism should be capable to account mentioned dependencies among the items in the negotiation set.

A negotiation strategy stands for the set of internal agent's rules it uses to pursue the goal of the encounter. Negotiation goal is often described in the terms of a deal stricken on the successful accomplishment of the encounter.

In our domain the goal of the meaning negotiation is to find the match between the contexts (negotiation set) of the background domain theories of the parties in order to align the domain ontology, and ensure that all parties commit to this common ontology in their subsequent interactions. A deal means a joint agreement and a commitment on the negotiated context and on the mapping of the concepts of the parties to the agreed concept.

In this paper we leave aside the discussion of the possible negotiation protocols and concentrate on the elaboration of the formal framework to design agents' strategies for multi-issue, non-symmetric meaning negotiation, provided that the agents play the aforementioned roles in a distributed information retrieval system.

3   Illustrative Example

Let's consider the following query submitted by a user agent *Q* to mediator agent *M*: "*Please retrieve all short papers by Ermolayev published in LNCS series*". The conceptual graph for this query is presented in Fig. 1a. We implicitly assume here that
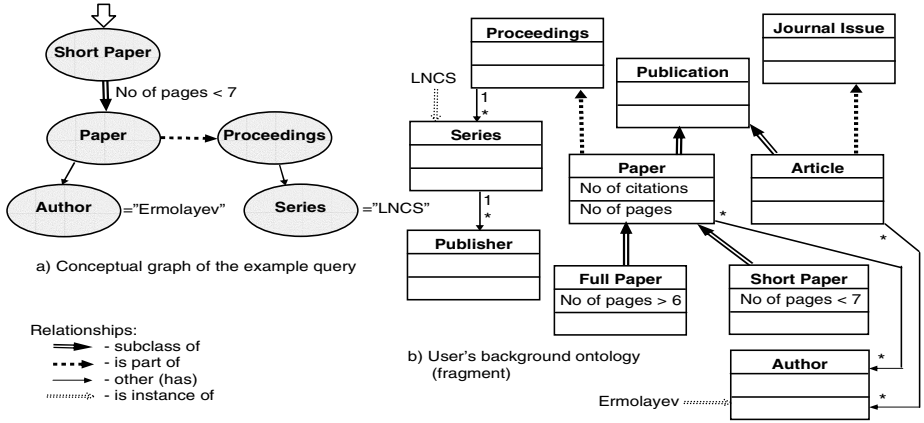
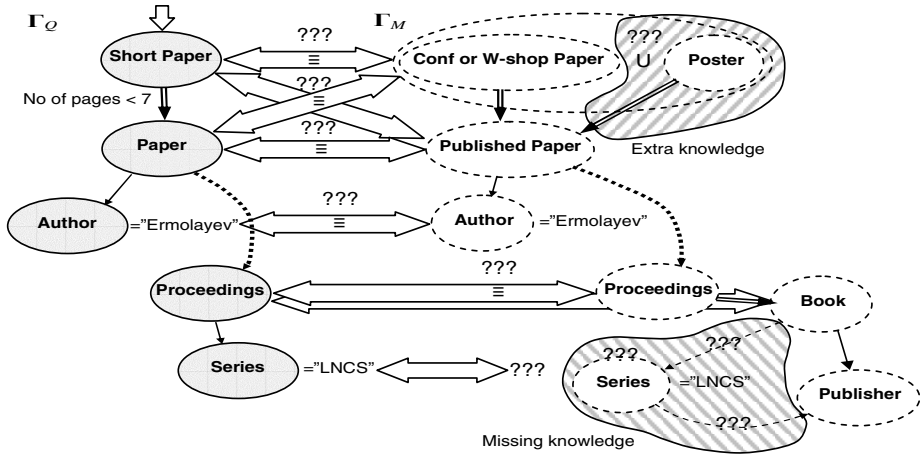**Fig. 1.** Example query and corresponding background domain theory

Fig. 2. Searching for the semantic match between Γ_Q and Γ_M

the corresponding background domain theory Γ_Q of Q is the one given in graphical notation in Fig. 1b. Following the approach of [7] we'll encode Γ_Q using the extended formalism of the Type Theory (TT) [8]:

$\Gamma_Q \vdash ShortPaper : \star_s$ (concept (type) *ShortPaper* is in the abstract type of sorts $\star_s$)

$\Gamma_Q \vdash Paper : \star_s$ (concept *Paper* is in the abstract type of sorts $\star_s$)

$\Gamma_Q \vdash Author : \star_s$ (concept *Author* is in the abstract type of sorts $\star_s$)

$\Gamma_Q \vdash Proceedings : \star_s$ (concept *Proceedings* is in the abstract type of sorts $\star_s$)

$\Gamma_Q \vdash Series : \star_s$ (concept *Proceedings* is in the abstract type of sorts $\star_s$)

$\Gamma_Q \vdash ShortPaper.NoOfPages : \star_a$ (property *NoOfPages* of the concept *ShortPaper* is in the abstract type of properties $\star_a$)

$\Gamma_Q \vdash ShortPaper < Paper : \star_s$ (concept *ShortPaper* is subsumed by the concept *Paper*)

$\Gamma_Q \vdash (N < 7) : ShortPaper.NoOfPages \rightarrow \star_p$ (proposition $(N < 7)$ applied to the property *NoOfPages* of the concept *ShortPaper* holds true – i.e. is in the abstract type of propositions $\star_p$)

$\Gamma_Q \vdash Datatype(N, Integer) : Paper.NoOfPages \rightarrow \star_p$ (proposition $Datatype(N, Integer)$
applied to the property $NoOfPages$ of the concept $Paper$ holds true – i.e. is in the abstract
type of propositions $\star_p$)
$\Gamma_Q \vdash has(P, Author) : Paper \rightarrow \star_p$ (as above)
$\Gamma_Q \vdash is\_part\_of(P, Proceedings) : Paper \rightarrow \star_p$ (as above)
$\Gamma_Q \vdash published\_in(R, Series) : Proceedings \rightarrow \star_p$ (as above)
$\Gamma_Q \vdash Ermolayev : Author$ (Instance $Ermolayev$ is of type $Author$)
$\Gamma_Q \vdash LNCS : Series$ (Instance $LNCS$ is of type $Series$)

This context of Γ_Q is submitted to M as the negotiation set. M will try to match it to
his domain theory Γ_M as graphically outlined in Fig. 2. The task for our paper is to
develop the formal framework for the strategy of negotiations between Q and M
which will allow them to automatically make this match as precise as possible. The
specificity of our settings is that:

- Background theories Γ_Q and Γ_M are not necessarily taxonomies – different types of
  semantic relationships should be accounted. This may imply various types of
  dependencies between the elements of the negotiation set.
- Background theories Γ_Q and Γ_M may be poor with instances. This may imply the
  necessity to use several kinds of semantic similarity metrices (not only instance
  similarity which proved to be quite precise [9] in comparing ontologies which are
  rich with instances).
- Background theories Γ_Q and Γ_M can not be disjoint in the sense that there will
  certainly be at least a partial match between them. The reason for this constraint is
  that normally a query is posed to the resources in the particular domain and the
  semantic discrepancies are at most surmountable.
- The cardinality of Γ_Q and Γ_M is moderate enough to allow NP-complete processing
  algorithms be acceptable. This actually means the constraint on Γ_M only because a
  query and its underlying domain theory are normally not bulky.

4   A Strategy for Meaning Negotiation

For simplicity reasons we'll further on consider that meaning negotiation occurs
between two parties (one to one negotiation) – a query submitter Q and a mediator M
agents. We shall build the framework for designing strategies for meaning
negotiations trying to provide a mechanism to find the context of a concept in Γ_M
closest to the context of a concept in Γ_Q of a query submitted by Q. As Γ_M and Γ_Q are
further on formalized as the sets of TT statements we shall build up our definition on
the notion of a context given in [7]:

Definition 1 (Context): The context $\mathbf{C}_c$ of a concept $c \in \Gamma$ is the union of the set Γ_i of
TT statements $\gamma_i \in \Gamma$ which are the assumptions over c and the set Γ_j of TT
statements $\gamma_j \in \Gamma$ which may be explicitly inferred from $\{\Gamma \vdash c : \star_s\} \bigcup \Gamma_i$ using the
rules of the type system [7]:

$$\mathbf{C}_c = \Gamma\big|_c = \Gamma_i \bigcup \Gamma_j. \tag{1}$$

To design such a negotiation strategy over Γ_Q we need to address the following problems:

– Which of the parties starts first?
– How to generate argumentation on the semantic discrepancies between Γ_Q and Γ_M?
– How to ensure that the ratio of these semantic discrepancies is monotonically decreasing in negotiation rounds?
– How to assess if the current level of these semantic discrepancies is sufficient to strike the deal?
– How to detect that the movement to the perfect match (no discrepancies) between Γ_Q and Γ_M is no longer possible?

Negotiations are evidently the series of mutually beneficial concessions. In the context of meaning negotiation we also need to denote what a concession is and how to compute the minimal effective concession.

Which of the parties starts first? The answer to this question in our settings is straightforward – the one who initiates negotiation by submitting a query makes the 1-st cry.

Argumentation on the semantic discrepancies and concession. We first need to denote how to formalize and to measure the semantic discrepancies between two contexts Γ_Q and Γ_M and then proceed with argumentation. It is natural to denote semantic discrepancies between two ontological contexts by means of the appropriately defined semantic distance **SD**, which is obviously a kind of a mapping $\mathbf{SD} : \Gamma_Q \times \Gamma_M \to \mathbf{R}$ (**R** is the set of real numbers). Suppose this mapping is defined (Section 5), then efficient argumentation should contain the set of presuppositions **PR** over Γ_Q and Γ_M which, if applied to Γ_Q, decreases **SD**. Some of the concepts, concept properties, or propositions expressing relationships from Γ_Q may have no analogy in Γ_M (or Γ_M in Γ_Q). We shall call these elements of Γ (Γ_Q or Γ_M) orphans. Local semantic distance SD_o between an orphan and Γ evidently has the maximal possible value. Argumentation on orphans should provide the counter-party with the information on the possible or anticipated context $\mathbf{C}_o$ to check it over his background domain theory and, possibly, find some extra context Δ_o to bridge the gap. If context Δ_o is found it extends the context of the party (Γ_Q or Γ_M). We shall say that a party concedes on the orphan o if $\mathbf{C}_o \cap \Delta_o \neq \varnothing$, i.e., if some of the arguments from $\mathbf{C}_o$ were accepted to Δ_o. Of course it is rational to make the smallest concession possible in a round because the acceptance of (the part of) $\mathbf{C}_o$ may induce the contradictions with another portions of Γ, which may result in harmonization overheads and, which is even more important, violate some of the collective commitments on the elements of Γ. From the other hand the concession should be sound enough for not to concede on the next negotiation round. Details are provided in Section 7.

How to make negotiation converging to a deal? We shall formulate the answer in the terms of the semantic distance **SD** between the contexts Γ_Q and Γ_M. Negotiation will converge to a deal if **SD** will be monotonically decreasing from round to round. A deal may be stricken between the parties if:

– No orphans are left in Γ
– The difference in **SD** in the current round and of the previous round equals to 0 or is less than the commonly agreed threshold

How to detect that further negotiation is useless? Negotiation is useless when all the parties have exhausted their argumentation and the deal is not stricken. In our settings this means that the orphans are still present in Γ, but there were no concessions in the two subsequent rounds. In this case the deal is impossible and the negotiation should result in failure. Practically this means that the party which submitted a query needs to reformulate it in the terms which are more coherent to the common ontology, or to give up.

5   Argumentation Using Propositional Substitutions and Contexts

Let $\mathbf{C}$ be the set of concepts in Γ_Q: $\mathbf{C} = \{c_i\}$. Evidently there may be several hypotheses on concept equivalence for each concept c_i. We'll denote the set of hypotheses on concept equivalence as: $\mathbf{H} = \bigcup_{i=1}^{n} H_i$, where n is the number of concepts defined in Γ_Q, H_i is the set of hypotheses on the equivalence of c_i to the concepts of Γ_M.
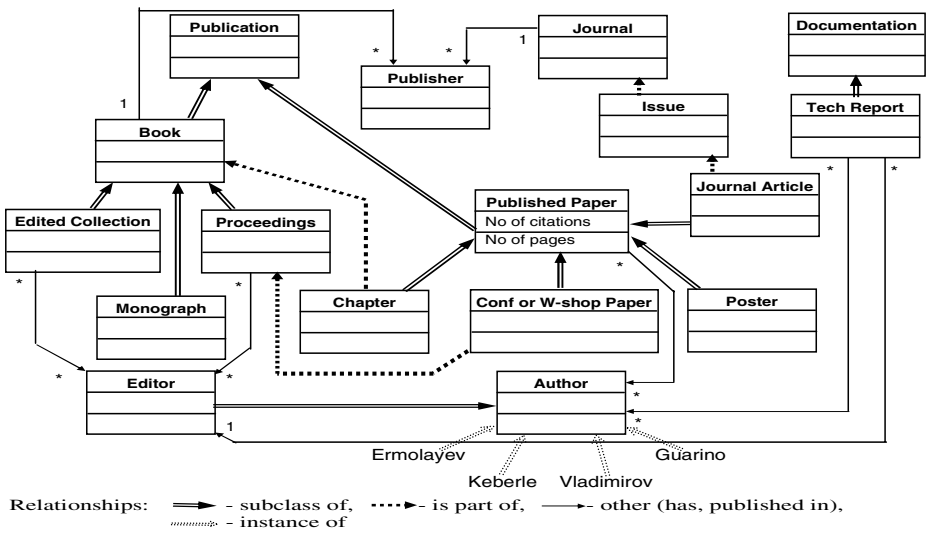

Fig. 3. Background domain theory Γ_M of M

Γ_Q provides a certain portion of facts on each c_i which is communicated by Q to M. For our example the portion of facts on the concept *ShortPaper* : $\star_s$ in TT notation is as follows:

$\Gamma_Q \vdash ShortPaper : \star_s$
$\Gamma_Q \vdash ShortPaper.NoOfPages : \star_a$
$\Gamma_Q \vdash ShortPaper < Paper : \star_s$
$\Gamma_Q \vdash (N < 7) : ShortPaper.NoOfPages \rightarrow \star_p$

M will try to apply this context to Γ_M (Fig. 3) and form hypotheses on the equivalence of c_i to the concepts of Γ_M. While forming these hypotheses M will exploit different applicable kinds of similarity measurement (Section 6) and, particularly, will try propositional substitutions for context similarity assessment.

Let:

- $\Gamma_Q\big|_{c_i} = \{\gamma_1,...,\gamma_n\}_Q\big|_{c_i}$ be the context of c_i in Γ_Q

- $h: c_{i_Q} \xleftrightarrow{\equiv} c_{k_M}$ be a hypothesis on the equivalence of c_i and c_k, $\Gamma_Q \vdash c_i$: $\bigstar_s$, $\Gamma_M \vdash c_k$: $\bigstar_s$

We'll say that γ'_j is the propositional substitution of $\gamma_j \in \Gamma_Q\big|_{c_i}$ if it is obtained by the substitution of c_i by c_k in γ_j. We'll say that the similarity asset of γ'_j to the evaluation of the context similarity s_j of γ'_j equals to 1 if $\Gamma_M \vdash \gamma'_j$, otherwise $s_j=0$.

We'll then compute the context similarity of c_i and c_k as $Sim_C = \dfrac{1}{m}\sum_{j=1}^{m} s_j$. If computed $Sim_C < 1$ M may provide Q with some context $\Gamma_M\big|_{c_k}$ to allow Q to make its hypotheses and assessments on the next negotiation round. The soundness of the hypothesis h may then be assessed by the overall similarity Sim_h of c_i and c_k. We'll compute Sim_h as the average of the measured similarities in frame of h.

For our example some of the hypotheses, contexts and similarity assessments generated by M are as follows:

Hypothesis: $Author_Q \xleftrightarrow{\equiv} Author_M$, **Assessment:** $Sim_h = 0.625$
Instance similarity:
Ermolayev : $Author_Q$
Ermolayev : $Author_M$; *Guarino* : $Author_M$; *Keberle* : $Author_M$; *Vladimirov* : $Author_M$;
$Author_Q \equiv Author_M$, $Sim_I = 0.25$
Lexical similarity:
$Author_Q \equiv Author_M$, $Sim_L = 1.0$

Hypothesis: $Proceedings_Q \xleftrightarrow{\equiv} Proceedings_M$, **Assessment:** $Sim_h = 1$
Lexical similarity:
$Proceedings_Q \equiv Proceedings_M$, $Sim_L = 1.0$

Hypothesis: $ShortPaper_Q \xleftrightarrow{\equiv} PublishedPaper_M$, **Assessment:** $Sim_h = 0.3$
Lexical similarity:
$ShortPaper_Q \equiv PublishedPaper_M$, $Sim_L = 0.3$
Propositional substitutions:
$\Gamma_M \vdash \neg((N < 7) : PublishedPaper_M.NoOfPages \rightarrow \bigstar_p)$, $Sim_h = 0$ – orphan constraint
$\Gamma_M \vdash \neg(PublishedPaper_M < Paper_Q : \bigstar_s)$, $Sim = 0$ – orphan subsumption
$\Gamma_M \vdash (Datatype(N, Integer) : PublishedPaper_M.NoOfPages \rightarrow \bigstar_p)$, $Sim = 1$
Context:
$\Gamma_M \vdash PublishedPaper_M < Publication_M : \bigstar_s$
$\Gamma_M \vdash JournalArticle_M < PublishedPaper_M : \bigstar_s$
$\Gamma_M \vdash Chapter_M < PublishedPaper_M : \bigstar_s$
$\Gamma_M \vdash ConfOrW\text{-}shopPaper_M < PublishedPaper_M : \bigstar_s$
$\Gamma_M \vdash Poster_M < PublishedPaper_M : \bigstar_s$
$\Gamma_M \vdash has(P, Author_M) : PublishedPaper_M \rightarrow \bigstar_p$
$\Gamma_M \vdash Datatype(N, Integer) : PublishedPaper_M.NoOfPages \rightarrow \bigstar_p$

Hypothesis: *ShortPaper$_Q$* $\xleftrightarrow{\equiv}$ *ConfOrW-shopPaper$_M$*, **Assessment:** *Sim$_h$*= 0.3
Lexical similarity:
ShortPaper$_Q$ $\equiv$ *ConfOrW-shopPaper$_M$, Sim$_L$*= 0,2
Propositional substitutions:
$\Gamma_M \vdash \neg ((N < 7) : ConfOrW\text{-}shopPaper_M.NoOfPages \rightarrow \star_p), Sim = 0$ – orphan constraint
$\Gamma_M \vdash \neg (ConfOrW\text{-}shopPaper_M < Paper_Q : \star_s), Sim = 0$ – orphan subsumption
$\Gamma_M \vdash (Datatype(N, Integer) : ConfOrW\text{-}shopPaper_M.NoOfPages \rightarrow \star_p), Sim = 1$
ShortPaper$_Q$ $\equiv$ *ConfOrW-shopPaper$_M$, Sim$_C$*= 0,3
Context:
$\Gamma_M \vdash ConfOrW\text{-}shopPaper_M < PublishedPaper_M : \star_s$
$\Gamma_M \vdash has(P, Author_M) : ConfOrW\text{-}shopPaper_M \rightarrow \star_p$
$\Gamma_M \vdash is\_part\_of(P, Proceedings_M) : ConfOrW\text{-}shopPaper_M \rightarrow \star_p$

Hypothesis: *ShortPaper$_Q$* $\xleftrightarrow{\equiv}$ *Poster$_M$*, **Assessment:** *Sim$_h$*= 0,6
Lexical similarity:
ShortPaper$_Q$ $\equiv$ *Poster$_M$, Sim$_L$*= 0
Propositional substitutions:
$\Gamma_M \vdash ((N < 7) : ConfOrW\text{-}shopPaper_M.NoOfPages \rightarrow \star_p), Sim = 1$
$\Gamma_M \vdash \neg (Poster_M < Paper_Q : \star_s), Sim = 0$ – orphan subsumption
$\Gamma_M \vdash (Datatype(N, Integer) : Poster_M.NoOfPages \rightarrow \star_p), Sim = 1$
ShortPaper$_Q$ $\equiv$ *Poster$_M$, Sim$_C$*= 0.6
Context:
$\Gamma_M \vdash Poster_M < PublishedPaper_M : \star_s$
$\Gamma_M \vdash has(P, Author_M) : Poster_M \rightarrow \star_p$
$\Gamma_M \vdash is\_part\_of(P, Proceedings_M) : Poster_M \rightarrow \star_p$

Hypothesis: *Paper$_Q$* $\xleftrightarrow{\equiv}$ *PublishedPaper$_M$*, **Assessment:** *Sim$_h$*= 0
Lexical similarity:
Paper$_Q$ $\equiv$ *PublishedPaper, Sim$_L$*= 0,5
Propositional substitutions:
$\Gamma_M \vdash \neg (is\_part\_of(P, Proceedings_Q) : PublishedPaper \rightarrow \star_p), Sim = 0$ – orphan meronymy
$\Gamma_M \vdash \neg (has(P, Author_Q) : PublishedPaper_M \rightarrow \star_p), Sim = 0$ – orphan relationship
Paper$_Q$ $\equiv$ *PublishedPaper$_M$, Sim$_C$*=0
Context:
$\Gamma_M \vdash PublishedPaper_M < Publication_M : \star_s$
$\Gamma_M \vdash JournalArticle_M < PublishedPaper_M : \star_s$
$\Gamma_M \vdash Chapter_M < PublishedPaper_M : \star_s$
$\Gamma_M \vdash ConfOrW\text{-}shopPaper_M < PublishedPaper_M : \star_s$
$\Gamma_M \vdash Poster_M < PublishedPaper_M : \star_s$
…

Hypothesis: *Series$_Q$* $\xleftrightarrow{\equiv}$ $\varnothing$, **Assessment:** *Sim$_h$*= 0 – missing knowledge in Γ_M
We shall measure the semantic distance D_i between c_i and Γ_M as follows:

$$SD_i = 1 - \max_{H_i}(Sim_{h_j}) . \tag{2}$$

We may now compute the overall semantic distance between Γ_Q and Γ_M as follows:

$$\mathbf{SD} = \sum_{i=1}^{n} SD_i . \tag{3}$$

6 Presuppositions

Extending the approach of [7] we'll make presuppositions on the equivalence of the concepts according to the measured Sim_h values. We'll then revise the propositional substitutions for other concept hypothesis and re-compute the corresponding Sim_h values. In result the presupposition set **PR** may be extended as well. Let $\mathbf{PR} = \bigcup_{i=1}^{n} PR_i$, where PR_i is the set of presuppositions on the equivalence of c_i. The rule for **PR** formation is as follows:

(1) Set up the similarity threshold *minSim* for accepting a hypothesis as the presupposition
(2) For each H_i:
− Choose the hypothesis h with the highest Sim_h value and add it to PR_i as *pr* iff its Sim_h value is over *minSim*
− Revise propositional substitutions for **H** with respect to *pr* and re-assess Sim_h values
(3) Repeat (2) until at least one *pr* is added to **H**
(4) For PR_i delete all *pr* except the one with the highest Sim_h value

After **PR** is formed we may also drop all the hypotheses in each H_i except the one with the highest Sim_h value. The difference in $\mathbf{SD}_b$ before and $\mathbf{SD}_a$ after the formation of **PR** will show us the efficiency of the formed **PR**: $E_{\mathbf{PR}} = (\mathbf{SD}_b\text{-}\mathbf{SD}_a)/\mathbf{SD}_b$.

For the presented fragment of our example $\mathbf{SD}_b = 2.775$. Presuppositions (*minSim*=0.5) with the highest Sim_h values are:

$Proceedings_Q \equiv Proceedings$, $Sim_h = 1$ and $Author_Q \equiv Author$, $Sim_h = 0.625$.

By revising propositional substitutions we obtain the following changes in *Sim* values:

$Paper_Q \xleftrightarrow{\equiv} PublishedPaper$, $Sim_h = 1$ (both propositional orphans are eliminated)
$Paper_Q \equiv PublishedPaper$, $Sim_h = 1$ may now be added to **PR**
$ShortPaper_Q \xleftrightarrow{\equiv} Poster$, $Sim_h = 1$ (subsumption orphan is eliminated)
$ShortPaper_Q \equiv Poster$, $Sim_h = 1$ may now be added to **PR**
$ShortPaper_Q \xleftrightarrow{\equiv} ConfOrW\text{-}shopPaper$, $Sim_h = 0.6$ (subsumption orphan is eliminated)
$ShortPaper_Q \equiv ConfOrW\text{-}shopPaper$, $Sim_h = 0.6$ may now be added to **PR**.
After **PR** is formed $\mathbf{SD}_a = 1.375$ and $E_{\mathbf{PR}} = 1.4/2.775 = 0.505$.

7 Concept Similarity

As it was mentioned before a negotiation set represented by the context of a background domain theory Γ_Q can not be treated as a well-defined rich ontology. Hence, we need to make all the efforts possible to assess its similarity to Γ_M (which is rather rich) through analyzing all facets of its semantics. We may achieve it only by following the advice of Weisberg [10]:

"... I would contend that analysts frequently should not seek a single measure and will never find a perfect measure. Different measures exist because there are different concepts to measure ... It is time to stop acting embarrassed about the supposed surplus of measures and instead make fullest possible use of their diversity."

The reminder of this section outlines the variety of similarity measures which we consider to be applicable for computing semantic distances in presuppositions. Most of these measures are widely used in ontology alignment [17].

Instance Similarity. The rationale behind the instance similarity is that similar concepts have similar instances. Let **D** be the domain, *A* and *B* be the concepts in **D**. *A* is similar to *B* if $I_A \cap I_B \neq \emptyset$, where I_A and I_B are the sets of instances of **D** and $I_A = \{i_k\} : \forall k, instance\_of(i_k, A)$. We'll follow the approach of [9] and measure Instance Similarity by symmetric Jaccard coefficient:

$$Sim_I(A,B) = \frac{P(I_A \bigcap I_B)}{P(I_A \bigcup I_B)} , \qquad (4)$$

where $P(I)$ is the probability that a randomly chosen instance of D belongs to *I*.

Context or Feature Similarity. The rationale behind the contextual similarity is that similar concepts have similar contexts. These contexts may be understood as feature sets. Similarity between feature sets may be measured for example by means of Tversky metrics [11]. However, Tversky metrics works only if there is a well defined and a commonly accepted feature set. This is not true in our case. Therefore we shall measure Context Similarity in frame of a hypothesis *h* through assessing the propositional substitutions as shown in Section 5: If s_i are the similarity assessments of the respective propositional substitutions γ'_j, then:

$$Sim_C = \frac{1}{m} \sum_{j=1}^{m} s_j . \qquad (5)$$

Datatype and Measurement Similarity. It seems rational to consider that similar concepts have similar properties. However, the problem of determining similarity among properties has the same complexity as measuring the similarity of concepts. Another observation is that the set of properties of a concept is the part of its feature set. Hence, it is worth trying to measure Property Similarity by a Context Similarity metrics. The peculiarity of a property is that there are different types of them: domain properties and referential (slot) properties. While a slot property is the matter reflecting the relationship to another concept (property), a domain property reflects that a concept has the feature which:

– Has a certain datatype (like a *colour*, a *weight*, an *age*, a *string*…)
– Is measured in certain (standard) units (like an *RGB vector*, a *kilogram*, a *year*, an *integer*, …)
– Has certain constraints on its values expressed as logical formulas, like: $(weight \leq 90) \wedge (age > 30)$

For example, if concept *A* (a *ShortPaper*) has property *a* having *integer* datatype and is measured in the *NoOfPages* and concept *B* (a *Poster*) has property *b* having *integer* datatype and is measured in the *NoOfPages*, *a* and *b* may be considered similar and, this fact may increase the similarity between *A* and *B*. The ratio of Sim_A in frame of the hypothesis $A \xleftrightarrow{\equiv} B$ will be increased if *a* and *b* have the same constraints. For example, $(a < 7)$ and $(b < 7)$ for a *ShortPaper* and a *Poster* respectively.

As far as Property similarity measurement is also based on the propositional substitutions we do not distinguish Sim_C and Sim_A in the discussion of the example in Section 5.

Lexical Similarity. Considering concepts with the same names (or the same lexical roots) may of course lead to confusion. However, the same root in names may be a good hint in finding a perfect match in a pair of concepts from different contexts like Γ_Q and Γ_M. Human experience says that this heuristics works if supported by other evidence (like instance similarity for $Author_Q$ and $Author_M$ or effective presuppositions based on the acceptance of $Proceedings_Q \xleftrightarrow{\equiv} Proceedings_M$ hypothesis in our example). We use the following lexical measure Sim_L. Let R_A, R_B be the sets of roots of the words which constitute the name of concept *A* and *B* respectively, then:

$$Sim_L = \frac{|R_A \cap R_B|}{|R_A \cup R_B|} .$$ (6)

In our example Sim_L of the concepts *Paper* and *PublishedPaper* equals to 0.5. Technically it is quite simple to build the sets of roots as far as the words which constitute a concept name often have capitalized first letters.

8   Concession and Reputation

What is concession with respect to a concepts' meaning in a multi-issue (i.e., multi-concept) negotiation? We'll denote this concession in the terms of concept similarity and satisfiability of the logical formulae describing the background theories of negotiating parties. Let: *M* be an actor in an information retrieval system, $\mathbb{N} = \{N_i\}$ be the set of its peers with whom *M* has agreed on the similarity of the concept A from its Γ_M to their concepts B_i from respective Γ_{N_i} in previous negotiations. We denote *Semantic Commitment* of *M* with respect to *A* as the set of hypotheses $H_A = \{h_{A,i} = (A \xleftrightarrow{\equiv} B_i, Sim_i)\}$ accepted by *M* as his beliefs. The strength of this *Semantic Commitment* may be assessed by:

$$S_A = \sum_i Sim_i .$$ (7)

When *Q* and *M* detect an orphan *A* in their negotiation a concession on *A* means the extension of Γ with $\mathbf{C}_A \cap \Delta_A \neq \varnothing$ by one of the parties (Section 4). Adopting this new portion of Γ for a party, say M, may force him to drop some of $h_{X,i}$ in his

Semantic Commitments with respect to the concepts *X* related to *A*. It is rational to consider that the party having less strong commitments (7) on the concepts related to *A* should concede.

A more knowledgeable party, having stronger *Semantic Commitments* may and is actually forced to concede less, i.e., require that the difference in similarity between its initial context and the agreed context is less than of the other party. Concessions of a more knowledgeable party will affect more commitments of the other parties which should of course be re-negotiated. It is therefore clear that negotiation strategy while determining the concession should:

– Avoid conceding on the concepts associated with strong *Semantic Commitments*
– If it is not possible to avoid concession, and there are alternative concepts to concede on – concede on the concept having *Semantic Commitment of* lower strength S_A

And, finally, the reputation of a party may be assessed by evaluating the extent to which its knowledge is really consensual. It is natural to consider a party more reputable if its overall strength of *Semantic Commitments* $\mathbf{S}^M = \sum_A S_A$ is greater than the $\mathbf{S}^N$ of another party. In our example it is evident that $\mathbf{S}^Q < \mathbf{S}^M$. Hence, it will be difficult for *Q* to convince *M* to accept the knowledge about *Series* concept (Fig. 2).

9 Discussion

The paper reports on the formal framework to design strategies for software agents engaged in multi-issue non-symmetric meaning negotiation. These agents are thought to be the actors in distributed information retrieval system based on centralized (e.g., [6], [13]) or de-centralized (e.g., [5]) mediator architecture with centralized, decentralized or hybrid ontology representation (please refer to [14] for a survey). In understanding the nature and the conceptual foundation of negotiation among software agents we base our work on the results from DAI and eCommerce domains. We adopt the theoretical basics of [2], [3] and build our conceptual framework on it. While designing a strategy for automated multi-issue meaning negotiation we address the aspects typical to a negotiation strategy in a more general sense (Section4). Speaking in terms of the Game Theoretical approach such a strategy should ideally lead to a deal in a reasonably small no of rounds and, if adopted by all negotiation parties, be in Nash equilibrium. Though it is not formally proved that the proposed strategy will be in the Nash equilibrium with similar strategy of another party, it seems to look like that according to the monotonic nature of similarity measures, the rules for forming presuppositions and concessions.

In the presented framework we used and extended various contributions of other authors in: measuring concept similarity [15], [9], [12], [11], using logical formulae to approximate semantic discrepancies [16], using Type Theory for formalizing argumentation [8], making presuppositions [7]. Sound experimental results of the colleges evaluating these basic contributions allow us to believe that our evaluation planned for the near future will bring positive results.

The advancements of the presented framework are as follows. It results in a strategy which compares the contexts of two background domain theories not concept by concept, but the whole context to the other context by accounting the relationships among concepts, the properties, the constraints over properties, and the available instances. The mechanisms for measuring contextual similarity through assessing propositional substitutions and to provide argumentation through generating extra contexts are also new. One more novelty of our framework is the use of presuppositions for choosing the best similarity hypotheses and to make concession to common sense monotonic. The means for evaluating the possible eagerness to concede through semantic commitments and related notions of knowledgeability and the degree of reputation to our knowledge also have not been reported before.

10   Concluding Remarks

As Tom Gruber said in one of his recent interviews:

> "I find it critical to remember that every ontology is a treaty – a social agreement – among people with some common motive in sharing."[1]

This view may definitely be applied also onto the artificial agents acting on behalf of their human owners. In this paper we have reported on the formal framework which provides strategies to gain such a social agreement automatically among software agents in an open system. These agents, as mentioned in Section 3, also have some common motive in sharing, though the semantics of their individual beliefs or background theories of the domain may well have surmountable discrepancies. We design multi-concept meaning negotiation strategies for software agents in distributed information retrieval which allow dealing on the common sense of a negotiated context and may be implemented in a software. Providing means to agree on the set of matters with respect to their semantics for autonomous intelligent components of an open software system is quite important. Such means will facilitate to making mutual understanding and collaborative work in such software systems more sound and effective.

Our plans for future work are to experiment with the prototype strategies based on the presented framework for multi-agent system in the successor of the RACING[2] project (distributed information retrieval). We also consider the implementation and the experimentation with such strategies in the frame of PSI[3] project. PSI objective is to prototype a multi-agent system which simulates dynamic engineering design processes and assists human designers in their cooperative work on a design project.

---

[1] Interview for the Official Quarterly Bulletin of AIS Special Interest Group on Semantic Web and Information Systems, Volume 1, Issue 3, 2004.

[2] RACING: Rational Agent Coalitions for Intelligent Information Retrieval on the Net, URL: http://www.zsu.zp.ua/racing.

[3] PSI: Productivity Simulation Initiative. Cadence Design Systems, GmbH.

References

1. Wiederhold, G.: Mediators in the Architecture of Future Information Systems. IEEE Computer, 25(3) (1992) 38–49
2. Lomuscio, R., Wooldridge, M., Jennings, N. R.: A Classification Scheme for Negotiation in Electronic Commerce" In: Dignum, F. and Sierra C. (eds.) Agent-Mediated Electronic Commerce: A European Perspective. Springer Verlag, Berlin Heidelberg, New York (2000), 19–33
3. Jennings, N. R., Faratin, P., Lomuscio, A. R., Parsons, S., Sierra, C., Wooldridge, M.: Automated Negotiation: Prospects, Methods and Challenges. Int. J. of Group Decision and Negotiation 10(2) (2001) 199-215
4. Beam, C., Segev, A.: Automated Negotiations: A Survey of the State of the Art. CITM Working Paper 96-WP-1022 (1997). URL: http://haas.berkeley.edu/~citm/wp-1022-summary.html
5. Zhang H., Bruce Kroft W., Levine B., Lesser V.: A Multi-agent Approach for Peer-to-Peer-based Information Retrieval Systems. In: Proc. AAMAS'04, July 19-23, 2004, New York, New York, USA
6. Ermolayev, V., Keberle, N., Kononenko, O., Plaksin, S., Terziyan, V.: Towards a framework for agent-enabled semantic web service composition. Int. J. of Web Services Research, 1(3) (2004) 63-87
7. Beun R.-J., van Eijk R.M., Prüst H.: Ontological Feedback in Multiagent Systems. In: Proc. AAMAS'04, July 19-23, 2004, New York, New York, USA
8. Luo Z.: Computation and Reasoning: A Type Theory for Computer Science. Int. Series of Monographs on Computer Science. Clarendon Press, Oxford (1994)
9. Doan, A., Madhavan, J., Domingos, P., Halevy, A.: Learning to Match Ontologies on the Semantic Web. Int. J. Very Large Data Bases, 12(4) (2003) 303-319
10. Weisberg, H.F.: American Political Science Review 68 (1974) 1638-1655
11. Tversky, A.: Features of Similarity. Psychological Review 84(4) (1977) 327-352
12. Bouquet, P., Magnini, B., Serafini, L., Zanobini, S.: A SAT-based Algorithm for Context Matching. Dept of Information and Communication Technology, Univ of Trento, T.R. # DIT-03-005, Jan. 2003
13. Bergamaschi, S., Castano, S., De Capitani di Vimercati, S., Montanari, S. Vincini, M.: An Intelligent Approach to Information Integration. In: Proc. Int. Conf. on Formal Ontology in Information Systems (FOIS-98), June, 1998.
14. Wache, H. et al.: Ontology-Based Integration of Information - A Survey of Existing Approaches. In: Gomez-Perez, A., Gruninger, M., Stuckenschmidt, H., Uschold, M. (eds.) Proc. of the IJCAI-01 Workshop on Ontologies and Information Sharing, Seattle, USA, August 4-5 (2001) 108-118
15. Lin, D.: An Information-Theoretic Definition of Similarity. In: Proc. Int. Conf. on Machine Learning (1998)
16. Aleksovski, Z., ten Kate, W., van Harmelen, F.: Semantic Coordination: a New Approximation Method and its Application in the Music Domain. In: Proc. ISWC-04 workshop on Meaning Coordination and Negotiation, 8 Nov.2004, Hiroshima, Japan (2004)
17. Euzenat, J. et al.: State of the Art on Ontology Alignment. KnowledgeWeb project deliverable D2.2.3, v.1.2. August 2, 2004. URL: http://knowledgeweb.semanticweb.org/