A Strategy for Automated Meaning Negotiation in Distributed Information Retrieval

Vadim Ermolayev
Natalya Keberle
Vladimir Vladimirov
Wolf-Ekkehard Matzke

Zaporozhye National University, Ukraine
Cadence Design Systems GmbH, Germany

Searching in the ISWC Semantic Bank: <negotiation> - One Item Found
The Outlook

• Motivation:
  – Google game or
  – Do we always use the PROPER domain theory?

• What happens in Distributed Information Retrieval:
  – Actors, Roles and the need to reach Agreements (on Domain Theories)

• Semantic Context and Negotiation Settings

• Meaning Negotiation Strategy:
  – How to behave smartly to reach agreements
  – Argumentation: Contexts, Propositional Substitutions, Presuppositions, Concession, Reputation ... and around

• Conclusions and future work

Shall be as informal as possible
Otherwise we’ll perish in endless deliberation
Do we Use the Proper Domain Theory?

• You work on agent-based system implementing a tourism-related application

• Who is inventing the same square wheel?

• One usual way to find out:
  – To ask a search engine:
    <agent> and <tourism> and <project>
  – E.g., Google:
    http://google.com/search?q=agent+tourism+project

• The results were …
Seems that ... we don’t – at all!

Links Found:  
- 141,000

Analyzed:  
- 1-50

Among them:  
Matches:  
- 13(26%)  
Mismatches:  
- 37(74%)
If We’ve been Smarter

• We should have used a **different** **DOMAIN** **ONTOLGY**

• This may have led us to …
... the Transformation Like:

• `<agent>` AND `<tourism>` AND `<project>`

  • **DOMAIN ONTOLOGY**

    `<agent>`: synonym_of(<agent>, <software agent>)
    `<software agent>`: is_a(<software agent>, <software>)
    `<software agent>`: implements(<software agent>, <recommender system>)
    `<software agent>`: component_of(<software agent>, <travel agent>)

  • `<tourism>` AND `<project>` AND `<software agent>` AND (recommender system OR <travel agent>)

• We have tried Google with that …
Is this the Proper Domain Theory?

Resources: 18
Among them:
Matches: 15 (+2 - 94%) vs 26% before
Mismatches: 3 (6%)

Interesting to note:
All of them could be found among the results (141 000) of the previous query

Compare: recall, precision
How to Adhere to the **PROPER** Domain Theory?

- Still not ready to answer
- We’ll explore what happens in **DIR** first …
Information Retrieval

Agents in Tourism Projects?
Information Retrieval
(semantically mediated – our Google game)

Software Agents as Recommender Systems in Tourism Projects?

Agents in Tourism Projects?
Distributed Information Retrieval
(agent-based, mediated... Semantic Context?)

Match? Align?
1 shot vs iterative

Semantic Context of a Query is too poor (incomplete) to provide reliable 1 shot matching
Distributed Information Retrieval
(agent-based, mediated, Negotiated Semantic Context)

RACING Mediator

- Query Formulation
- Query Transformation
- Query Decomposition
- Negotiation with IRPAs
- Sub-Query Outsourcing

IRP (Agent)
How to Adhere to the **PROPER** Domain Theory?

- Just observe **what people do:**
  - Be **smart**
  - Don’t be **stubborn**
  - Be ready to **concede**
    - As much as your reputation allows
  - Be **pro-active**
    - Try to reach the agreement on the Semantic Context of the Query

- **Negotiation** - incorporating all of the **above**
  - Use **Argumentation** to negotiate
  - In a way to **Concede monotonically** to the **Deal**
Negotiation Settings:
One-to-One, Non-Symmetric, Multi-Issue, on Semantic Context

• The **Goal**
  – The **Deal** stricken over the **Negotiation Set**

• The Interaction **Protocol**
  – Symmetric vs **Non-Symmetric**
  – **One-to-One**, **One-to-Many**, **Many-to-Many**

• The **Negotiation Set**
  – Single-Issue vs **Multi-Issue**
  – **Semantic Context** (the part of the Domain Theory communicated to the negotiation party)

• The **Strategy** (of a party)
  – The set of internal Rules an Agent uses to pursue the **Goal** (of striking the Deal)
Semantic Context
after (Beun, van Eijk, and Prüst, 2004)

• **Definition 1** (Semantic Context): The context $C_c$ of a concept $c \in \Gamma^*$ is the union of the set $\Gamma_i$ of $\mathbb{T}\mathbb{T}$ statements $\gamma_i \in \Gamma$ which are the assumptions over $c$ and the set $\Gamma_j$ of $\mathbb{T}\mathbb{T}$ statements $\gamma_j \in \Gamma$ which may be explicitly inferred from $\{\Gamma \vdash c : \star_s \} \cup \Gamma_i$ using the rules of the type system:

$$C_c = \Gamma|_c = \Gamma_i \cup \Gamma_j$$

* $\Gamma$ stands for **Domain Theory**
** $\mathbb{T}\mathbb{T}$ stands for **Type Theory**


*** $\Gamma \vdash c : \star_s$ reflects that 1) $c$ is the concept (has the special type “sort”) and 2) this fact (1) may be inferred from the Domain Theory
Negotiation Strategy: the Questions to be Answered (by providing the Rules)

• Let $Q$ has $\Gamma_Q$ and $M$ has $\Gamma_M$:
  – Which of the parties starts first? – Straightforward! $Q$ of course

• The others are more difficult:
  – How to generate argumentation on the semantic discrepancies between $\Gamma_Q$ and $\Gamma_M$?
  – How to ensure that these discrepancies are eliminated monotonically in negotiation rounds?
  – How to assess if the current level of these semantic discrepancies is sufficient to strike the deal?
  – How to find out that the movement to the perfect match (no discrepancies) is no longer possible?
Argumentation on Semantic Discrepancies

• Define **Semantic Distance** as $SD : \Gamma_Q \times \Gamma_M \rightarrow R$

• Efficient argumentation should lower the $SD$ (monotonically)

• Biggest contribution to $SD$ is provided by the “**orphans**” of $\Gamma_Q$ wrt $\Gamma_M$ (or $\Gamma_M$ wrt $\Gamma_Q$)
  – **Orphans**: concepts, concept properties, or propositions expressing relationships of $\Gamma_Q$ having no analogy in $\Gamma_M$ (or of $\Gamma_M$ in $\Gamma_Q$)

• So – find a kind of an extra context $\Delta_o$ for each encountered orphan, say, $o$

• A party concedes on $o$ if $C_o \cap \Delta_o \neq \emptyset$

Orphans: an Example
The Google Game

One can find a different (more detailed) example in the paper
Contexts & Propositional Substitutions

• **Q -> the Context of a Project:**
  - An Agent implements a Project

• **M -> Equivalence hypotheses:**
  - Agent$_Q$ ↔ Agent$_M$
  - Agent$_Q$ ↔ Software Agent$_M$

• **M -> Propositional substitution:**
  - Software Agent implements a Recommender System

• Communicated to **Q** as the Argumentation (Context)

• By making Presuppositions
Presuppositions

• Based on the computed Sim values
• M - Presupposition: Project_Q ↔ Recommender System_M
• M: What if Q submitted
  – An Agent implements a Recommender System
• But NOT
  – An Agent implements a Project
• The Sim value of Agent_Q ↔ Software Agent_M will GROW
• Formally: Presupposition Set PR = \bigcup_{i=1}^{n} PR_i is formed wrt the communicated context C
Presuppositions Make Contexts Closer
Presuppositions Make Contexts Closer
Presuppositions Make Contexts Closer
Presuppositions Make Contexts Closer
A Presupposition becomes the Propositional Substitution

\[ h: \text{Project}_Q \text{ equals to } \text{Recommender System}_M \]
The Use of Presuppositions

• (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 the propositional substitutions for $H$ wrt $pr$ and re-assess $Sim_h$ values
• (3) Repeat (2) until at least one $pr$ is added to $H$
• (4) For $PR_i$ drop 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 $SD_b$ before and $SD_a$ after the formation of $PR$ shows the efficiency of the formed $PR$:
  $$E_{PR} = (SD_b - SD_a) / SD_b$$
When to Stop?

• A deal may be stricken if:
  – No orphans are left in $\Gamma_Q$ wrt $\Gamma_M$ (or $\Gamma_M$ wrt $\Gamma_Q$)
  – Some orphans are still present, but SD is less than the commonly agreed threshold

• Further negotiation is useless (the parties have exhausted their argumentation and end up without the deal):
  – The (substantial) orphans are still present
  – There were no concessions in the two subsequent rounds
  – Q needs to reformulate the query it in the terms more coherent to $\Gamma_M$ or to give up
More Semantic Commitments –
Less Freedom to Concede

• The encounter is **non-symmetric**

• **M** normally has lots of **Semantic Commitments** to keep (agreements on similarities or even equivalence)

• **Q** may offer a good reason to drop some of them
  – If **M adopts** – than needs to re-negotiate with all the others (lots of risk that some peers abstain)
  – If **M abstains** – no concession – risk to end up with no deal (locally)

• So **M** will better **abstain**

• The **Readiness to Concede** should be **weighted by** the degree of the **Semantic Commitment** of the party:
  – **Q** should be ready to **concede more** (to receive the service)
  – **M**’s **reputation** makes it **more stubborn**
Conclusions and Future Work

• We are at an early stage
• The formal framework has been developed in RACING*
• Partly adopted by PSI* Negotiation Framework
• Ontology debate framework (1 PhD student working)
• Research Prototype implementation anticipated
• Evaluation experiments
  – E.g., like the extended Google game …
  – As one of the reviewers wrote – a challenging task itself …
• Looking forward to receiving advice
• Ready for cooperation

* Please ask for back-up slides
“I find it critical to remember that every ontology is a treaty – a social agreement – among people with some common motive in sharing.”

- Tom Gruber (recently)

Propositional Substitution: **People <- Agents**

Questions please ...
• **Title:** Rational Agent Coalitions for INtelliGent Mediation of Information Retrieval on the Net

• **Objective:**
  – Investigate and evaluate the applicability of agent-based approach covering rationality, agency, coalition formation, collaboration to market oriented sectors of Distributed Information Retrieval

• **Focus:**
  – Mediation of information search and retrieval from structured or weakly structured information resources of:
    – Full-text online collections of Scientific Publications
    – Online Teaching Materials

• **Performed by:**
  – Dept of IT, Zaporozhye National University

• **Funded by:**
  – Ukrainian National Ministry of Education and Science

• **URL:**
Productivity Simulation Initiative
Project Lines and Partners

Governmental Funding

Industrial Sponsors

2004
MF
SSP
NF
QF
AF
UC
TBC
TBC
TBC
Industrial Product
UP
PPE
PAS
CERTICON
Beyond …

Research
Evaluation