Strategies for Ontology Negotiation: Finding the Right Level of Generality

Jurriaan van Diggelen, Edwin D. de Jong, Marco A. Wiering
Institute of Information and Computing Sciences
Utrecht University, the Netherlands
{jurriaan, dejong, marco}@cs.uu.nl

ABSTRACT

In heterogeneous multi agent systems, communication is hampered by the lack of shared ontologies. Ontology negotiation is a technique that enables pairs of agents to overcome these difficulties by exchanging parts of their ontologies. As a result of these micro level solutions, a communication vocabulary emerges on a macro level. The goal of this paper is to ensure that this communication vocabulary contains words of the right level of generality, i.e. not overspecific and not overgeneralized. We will propose a number of communication strategies that enable the agents to achieve these goals. Using experimental results, we will compare their performance.

1. INTRODUCTION

A fundamental communication problem in open multi agent systems (MAS's) is caused by the heterogeneity of the agent's knowledge sources, or more specifically of the underlying *ontologies*. Although ontologies are often advocated as a complete solution for knowledge sharing between agents, this is only true when all agents have knowledge about each others' ontology. The most straightforward way to establish this would be to develop one common ontology which is used by all agents [9]. However, this scenario would be very unlikely in open multi agent systems, as those on the internet, because it would require all involved system developers to reach consensus on which ontology to use. Moreover, a common ontology forces an agent to abandon its own world view and adopt one that is not specifically designed for its task [3]. This may result in a suboptimal situation.

Ontology negotiation [2] has been proposed as a technique that enables agents to preserve their local ontologies, and solve communication problems at agent-interaction time. Communication problems between heterogeneous agents are solved by establishing a shared communication vocabulary (or CV). Communication proceeds by translating from the speaker's local ontology to the communication vocabulary, which the hearer translates back to its own local ontology. When two agents start communicating, they first try to cope with the situation as is. When the speaker uses a word that the hearer does not understand, it solves the problem at hand by teaching the meaning of this word to the hearer. This enables two agents that regularly communicate with each other to build towards a solution for their semantic integration problem on an as-need basis.

Whereas an ontology negotiation protocol provides a nice solution to incrementally establish a communication vocabulary between a *pair* of heterogeneous agents, it is not straightforward how this solution scales to *whole* multi agent systems. A decentralized approach such as ontology negotiation may give rise to a proliferation of different CV's between different agent pairs in the system. This would be disadvantageous for the agents, as agents would have to use different words with different agents, which would make communication unnecessarily complicated. Furthermore, agents would have to spend much effort on building CV's, as the CV that has been built up with one agent may not be useful for communication with another agent. Therefore, when two agents participate in ontology negotiation to resolve their mutual misunderstandings, they should also pursue the goal of establishing a uniform and effective CV for the benefit of the whole community.

In this paper we will describe communication strategies for ontology negotiation protocols that take this global goal into account. These strategies prescribe which words and meanings the agents should teach each other during ontology negotiation. Regarding the words, we aim for a situation where every agent uses the same unique word for the same meaning. This is to be established by the agent's *word selection strategy* which we have studied in earlier work [5]. Regarding the meanings, we aim for a communication vocabulary which enables the agents to communicate at the right level of generality. Agents with different areas of expertise should not communicate at an *overspecific* level, as not everything that is of interest to one agent is also of interest to another agent. To prevent the CV from becoming bulky and difficult to learn, the CV should not contain such overspecific meanings. However, the meanings in the CV should not be *overgeneralized* either to enable the agents to convey sufficient information. Finding the right balance between specificity and generality of words is to be established by the *meaning selection strategy*. In this paper, we will show how a well designed meaning selection strategy contributes to faster semantic integration in the group of agents.

In the next section, we review related work. In Section

.

3, we describe the framework and explain how the communication protocols and strategies fit in. Section 4 presents the model that is used for the experiments. According to that model, some integration measures are proposed that measure the degree of semantic integration. Section 5 gives a precise description of the meaning selection strategy. In Section 6 the results of the experiments are presented, and the different meaning selection strategies are compared. We conclude in Section 7.

2. RELATED WORK

Most solutions that have been proposed for semantic integration problems are not flexible enough to be suitable for large open MAS's. Approaches such as ontology alignment [12] require ontologies to be aligned before the agents start interacting. In open MAS's it is not known beforehand which agents will interact with each other, and therefore, one can not tell in advance which ontologies must be aligned.

Ontology agents [1, 13] and *mediation services* [11] have been proposed as central services that reconcile heterogeneous ontologies at agent-interaction time by translating between ontologies. Such services have access to a library of concept-mappings between every ontology in the system. In large open MAS's, such a library would become too complex to be reliably maintainable.

Therefore, for large open systems, a *decentralized* technique is needed that allows agents to solve ontology problems among themselves at the time they arise. W. Truszkowski and S. Bailin have coined the term *Ontology Negotiation* to refer to such approaches [2]. Other approaches for ontology negotiation are [18, 17, 7].

Because the field of ontology negotiation is relatively new, and it is a very ambitious approach to achieve semantic integration [16], there are still many open problems. One of these problems is how a uniform CV that is shared among the whole *group of agents* may result from conversations that have taken place between *pairs of agents*. The question how a global language system arises from the interactions between individual agents is well studied in the language evolution community [15, 4]. Most of these approaches serve an explanatory goal, i.e. understanding how a communication system may evolve in a group of heterogeneous agents. Our goals, however, are purely constructive, i.e. we aim at designing communication strategies that can be used during ontology negotiation in order to establish a communication vocabulary of a certain quality. In particular we aim for an *optimal distributed communication vocabulary* [6], meaning that the CV is minimal in size and sufficiently expressive. One of the ways to make the CV minimal in size, is to ensure that the agents communicate at the right level of generality, which is the topic of this paper.

3. FRAMEWORK

3.1 Ontologies and vocabularies

Figure 1 shows an example of two agents in our framework. The dashed rectangle shows the meaning space in the system, i.e. the meanings that are assumed to exist in the environment of the agents. In the example, the meanings that constitute the meaning space are *m1* to *m8*.

The agent's ontology assigns names to meanings in the meaning space. For example, the ontology of Ag1 specifies

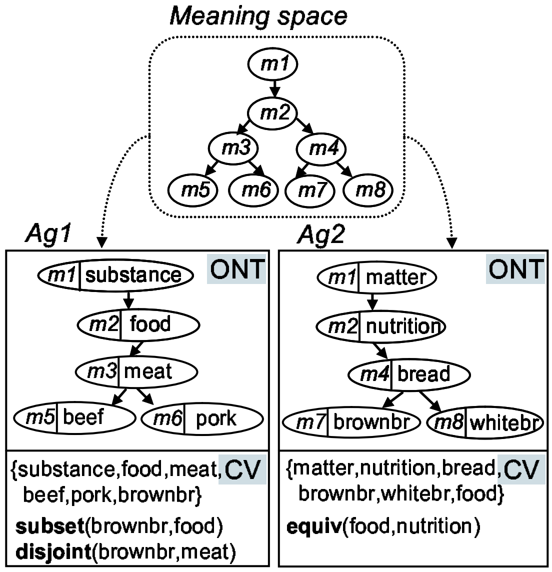

Figure 1: Example ontologies

that *m1* is called "substance" and that *m2* is called "food". A meaning with its corresponding name will be called a *concept*. An arrow from concept c to concept d represents that concept c is more general than concept d, and conversely that concept d is more specific than concept c.

Not every agent assigns the same names to the meanings in the meaning space. For example, Ag1 calls *m1* "substance" and *m2* "food", whereas Ag2 calls *m1* "matter" and *m2* "nutrition". To avoid naming-conflicts (two agents assigning the same name to different meanings), we assume that every agent uses a unique set of names in its ontology. This can be easily achieved by prefixing the names in the ontology with namespaces.

The property that every agent in the system uses distinct names to represent meanings is one source of the heterogeneity of the ontologies. Another source is that the ontologies of the agents contain concepts that correspond to different meanings. For example, the meanings *m4*, *m7* and *m8* are present in the ontology of Ag2, but are not present in the ontology of Ag1. This is a typical characteristic of heterogeneous multi agent systems, where every agent uses an ontology that is tailored to its own specific task. For example, Ag1 can be thought of as being a butcher as its ontology reflects expertise on meat. Ag2 can be thought of as being a baker as its ontology reflects expertise on bread.

Whereas the agents use the concepts in their ontology (ONT) for local knowledge representation and reasoning, for communication they use their communication vocabulary (CV). Note that the words in the communication vocabulary are not necessarily shared with the other agents. The CV contains the words that an agent may use to communicate something, regardless whether this word will actually be understood by the listener or not. Initially, the CV of an agent contains only the names of the concepts in its local ontology, as these are the only words that it knows for the meanings in its ontology. Because these words are unique, none of the other agents will understand them. When an agent is not understood by another agent, it explains the meaning of the uncomprehended word, after which the listener adds the word to its communication vocabulary. For example, the

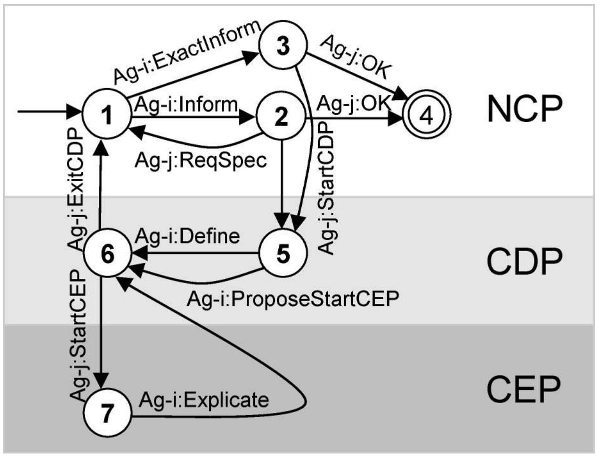

Figure 2: Message protocol

word "food" in the communication vocabulary of Ag2 is the result of a conversation in which Ag1 used "food", Ag2 did not understand it, after which Ag1 taught the meaning of "food" to Ag2. This teaching process enabled Ag2 to formulate a definition of the word "food" in terms of its ontology, namely "**equiv**(food,nutrition)". This states that the word "food" is equivalent in meaning with the "nutrition" concept in its ontology. A definition may also state that a word in the CV means something more specific than a concept in the ontology. For example, "**subset**(brownbr,food)" in the CV of Ag1.

3.2 Communication protocol

Figure 2 shows an ontology negotiation protocol that is used in ANEMONE [7]. Using this protocol, agents like Ag1 and Ag2 may successfully communicate if this is enabled by their communication vocabularies. Otherwise the agents extend their communication vocabularies to make communication possible. Three layers can be distinguished in this protocol. The upper layer is the Normal Communication Protocol (NCP), which deals with information exchange between the agents. If this is not possible, the agents switch to the Concept Definition Protocol (CDP), where the agents give a definition of a word in terms of other words. If this is not possible (when the listener does not understand the definition), the agents switch to the Concept Explication Protocol (CEP), where the agents convey the meaning of a word by pointing to examples.

We will explain the protocol in further depth below. Communication starts in state 1 where Ag-i wishes to communicate a meaning from its ontology to Ag-j. For example, suppose that Ag2 wishes to communicate the meaning *m8* (corresponding to the concept "whitebr") to Ag1. In state 1, Ag2 must select an appropriate word in the communication vocabulary to communicate *m8*. There are different possibilities for this. The first possibility is to select a word in the CV that is equivalent in meaning with *m8* (such as the word "whitebr"), and send a message "ExactInform(whitebr)" after which it ends up in state 3. If the CV of Ag1 would have contained the word "whitebr", Ag1 would have translated this word to its own ontology, and responded "OK". As Ag1 does not know the meaning of "whitebr", it responds with "StartCDP" to incite Ag2 to convey the meaning of "whitebr" in the Concept Definition Protocol. Another possibility for Ag2 to convey the meaning *m8* is to choose a word in the CV that means something more general than *m8* (such as the words "bread", "nutrition", "food" or "matter"), and send a message with "inform" after which it ends up in state 2. When Ag1 does not know the word used in the message, it responds "StartCDP" to start the Concept Definition Protocol. If Ag1 knows the meaning of the word, it checks whether the message is not overgeneralized. If it believes the message might be overgeneralized, it responds "ReqSpec" (Request specification) to incite Ag2 to use a more specific word. If Ag1 assesses that the message is not overgeneralized, it translates the message to its ontology and responds "OK". The method for recognizing overgeneralized messages we use here is a simplified version of the one used in the ANEMONE protocol. If the receiver's ontology contains no concepts that mean something more specific than the word in the message, the receiver assesses that the message is not overgeneralized. In this case, the receiver regards requesting for a more specific word useless, because its ontology is not fine grained enough to process any extra information. If the receiver's ontology contains concepts that are more specific than the meaning of the word, the receiver believes that the message might be overgeneralized and responds "ReqSpec".

The agents enter the Concept Definition Protocol in state 5, where Ag-i defines the meaning of the word in terms of other words in the communication vocabulary. Suppose that Ag2 wishes to define "nutrition", it sends a message "Define(equiv(nutrition, food))" to Ag-1, which enables Ag1 to derive the definition of "nutrition" after which Ag1 answers "ExitCDP". If the receiver of the "Define" message does not understand the definition of a message, it responds "StartCEP" to start the Concept Explication Protocol which incites the sender to explicate the meaning of the word by pointing to examples. If the sender of the definition is not able to give a definition (for example, Ag2 does not know any other word for "whitebr"), it sends the message "ProposeStartCEP".

In the Concept Explication Protocol (state 7), the agent conveys the meaning of the word by giving a set of positive and negative examples. More information on this type of concept learning can be found in [8].

3.3 Communication strategies

Having described the ontology negotiation protocol, we will now describe how the communication strategy fits in.

Word selection strategy

Suppose that Ag2 has the intention to convey the meaning *m2*. It has two words in its communication vocabulary that correspond to this meaning, namely "nutrition" and "food". The *word selection strategy* selects one of these word. In previous work [5], we have shown that the most effective word selection strategy is to choose the word that has most frequently been used by other agents. In this paper, we will use this word selection strategy, and focus on the other communication strategy: the meaning selection strategy.

Meaning selection strategy

Consider again the situation in state 1 of the protocol where Ag2 intends to convey the meaning *m8* ("whitebr"). As has been argued in the previous section, Ag2 may convey this meaning by choosing a word that means *m8* or a word that means something more general than *m8*, i.e. a word that means *m4*, *m2* or *m1*. The meaning selection strategy prescribes which meaning Ag2 should choose. A good meaning selection strategy selects a meaning that is not overgeneralized in order not to provoke the response "ReqSpec". However, the meaning selection strategy should not select a meaning that is too specific either, to prevent the communication vocabulary from becoming large and filled with words that are unnecessarily specific. Examples of overgeneralized concepts are *m1* and *m2*, as from a god's eye perspective we can predict that this will provoke a "ReqSpec" answer from Ag1. An example of an overspecific concept is *m8*, because from a god's eye perspective we can determine that this word contains superfluous information for Ag1. *m4* is at the right level of generality. It is not overspecific as it is more general than *m8* and thereby more widely applicable. Furthermore, from a god's eye perspective, we can assess that it is not overgeneralized as it will not provoke a "ReqSpec" answer.

Of course, the agents do not have access to this god's eye perspective. They therefore do not know which words are overgeneralized and which are overspecific. The difficulty of the meaning selection strategy lies in the making of an *educated guess* which word is at the right level of specificity. Before we describe how this can be done, we will present the model in which we can test different strategies.

4. MODEL

The experiments are performed using a set of agents $MAS = \{Ag_1..Ag_n\}$. The ontologies of the agents are randomly created and, like the ontologies in Figure 1, may cover different parts of the meaning space. The formal counterpart of the meaning space in Figure 1 is defined using graph theory [10]. A meaning space M is defined as a rooted tree (V, E), where V is a set of vertices, E is a set of directed edges, and a particular vertex in V is designated as the root. A vertex v_j is a child of vertex v_i iff $\langle v_i, v_j\rangle \in E$. A vertex with no children is called a *leaf*; a vertex that is not a leaf is called *internal*. A vertex v_j is a *descendant* of vertex v_i (and conversely v_i is an *ancestor* of v_j) iff there is a directed path from v_i to v_j. If T is a rooted tree with root v_0, then $ln(v_i)$ denotes the *level number* of v_i which equals the length of the unique directed path from v_0 to v_i. The depth of a tree is the largest level number achieved by a vertex in that tree. The following definition is useful to characterize the shape of a meaning space.

DEFINITION 1. *A meaning space $M = (V, E)$ is defined according to $B = (b_0, .., b_d)$ if:*

- *d is the depth of the tree M*
- *for each $v_i \in V$, v_i has $b_{ln(v_i)}$ children*

For example, the meaning space in Figure 1 is defined according to (1,2,2,0), because $m1$ (at level number 0) has 1 child; $m2$ (at level number 1) has 2 children; $m3$ and $m4$ (at level number 2) have 2 children; $m5$, $m6$, $m7$ and $m8$ (at level number 3) have 0 children.
An ontology ONT is defined as a tuple $\langle \mathcal{C}, M, \mathcal{I}\rangle$, where $\mathcal{C}$ is a set of concept names, $M = (V, E)$ is a meaning space and $\mathcal{I}$ is a bijective mapping from $\mathcal{C}$ to V. To be able to characterize the ontologies in the system, we use the following definition

DEFINITION 2. *Given an ontology $ONT = \langle \mathcal{C}, M, \mathcal{I}\rangle$, where $M = (V, E)$. ONT is defined according to B and B_g if*

- *M is defined according to B, and*
- *$V \subseteq V'$, $E \subseteq E'$, where*
  - *$M' = (V', E')$ is a meaning space defined according to B_g.*

For example, the ontologies of Ag1 and Ag2 in Figure 1 are defined according to $B = (1, 1, 2, 0)$ and $B_g = (1, 2, 2, 0)$.

4.1 Integration Measures

In this section, we will define some measures which indicate how well the agents can understand each other. Suppose that Ag_i wishes to communicate a meaning m to Ag_j. If Ag_i can do this in only the NCP layer (the upper layer in the protocol of Figure 2), the understandings rate between Ag_i and Ag_j with respect to meaning m is 1; if the agents have to visit the CDP or CEP layer, the understandings rate is 0.

DEFINITION 3. *MPUR: Meaning and Pair dependent Understandings Rate.*
$MPUR(m, \langle Ag_i, Ag_j\rangle)$ is

- *1 if the conversation to communicate m from Ag_i to Ag_j finishes without visiting the CDP and CEP layer*
- *else 0.*

The following measure indicates how well an agent Ag_i can communicate an average concept to Ag_j (ONT_i is defined as a tuple $\langle \mathcal{C}, \langle V, E\rangle, \mathcal{I}\rangle$, according to definition 2):

DEFINITION 4. *PUR: Pair dependent understandings rate*
$PUR(\langle Ag_i, Ag_j\rangle) = \frac{1}{\#V_i} \sum_{m\in V_i} MPUR(m, \langle Ag_i, Ag_j\rangle)$

The following measure indicates how well an average agent can communicate an average meaning to an average other agent.

DEFINITION 5. *UR: Understandings rate*
$UR = \frac{1}{n^2} \sum_{Ag_i, Ag_j \in MAS} PUR(\langle Ag_i, Ag_j\rangle)$

If the understandings rate is 1, every agent can communicate everything to every other agent.

5. FINDING THE RIGHT LEVEL OF GENERALITY

Using the different integration measures introduced in the previous section, we can characterize overgeneralized and overspecific concepts in further depth.

5.1 From a god's eye view

PROPERTY 1. *Teaching overgeneralized concepts does not increase MPUR (definition 3).*

We will illustrate this property using the example where Ag2 intends to communicate *m8* (the meaning of "whitebr") to Ag1. Suppose Ag2's meaning selection strategy selects the overgeneralized meaning *m1* (the meaning of "matter"). Before Ag2 sends this message, $MPUR(m8, \langle Ag2, Ag1\rangle) = 0$

(because Ag1 does not understand the word "matter"). After Ag2 has taught the concept "matter" to Ag1, $MPUR(m8, \langle Ag2, Ag1\rangle)$ still equals 0 (because "matter" invokes a "ReqSpec" response and Ag2's second attempt to convey *m8* fails). Now suppose that Ag2's meaning selection strategy selects the meaning *m4* (corresponding to the word "bread"). This meaning is not overgeneralized, because $MPUR(m8, \langle Ag2, Ag1\rangle)$ becomes 1 after the concept "bread" has been taught to Ag1 (because "bread" invokes an "OK" response).

PROPERTY 2. *Teaching overspecific concepts gives rise to little increase in PUR (definition 4).*

Consider again the situation where Ag2 intends to communicate *m8* ("whitebr") to Ag1. Suppose that the CV's of Ag1 and Ag2 are still in their initial configuration, i.e. they only contain the names of the concepts in their ontologies. Suppose that Ag2's meaning selection strategy selects the meaning *m8* (corresponding to the word "whitebr"). Before Ag2 sends this message, $PUR(\langle Ag2, Ag1\rangle) = 0$ (Ag2 can not communicate anything to Ag1). After Ag2 has taught the word "whitebr" to Ag1, $PUR(\langle Ag2, Ag1\rangle) = \frac{1}{5} \cdot MPUR(m8, \langle Ag2, Ag1\rangle) = \frac{1}{5}$. Now, suppose that Ag2's meaning selection strategy would have selected "bread". After Ag2 has taught the word "bread" to Ag1, $PUR(\langle Ag2, Ag1\rangle) = \frac{1}{5} \cdot (MPUR(m8, \langle Ag2, Ag1\rangle) + \frac{1}{5} \cdot MPUR(m7, \langle Ag2, Ag1\rangle) + \frac{1}{5} \cdot MPUR(m4, \langle Ag2, Ag1\rangle)) = \frac{3}{5}$. Note that, compared to the word "bread", the teaching of the word "whitebr" gives rise to little increase in understandings rate between the pair (and therefore also in understandings rate in general). This is why "whitebr" is overspecific, and "bread" is not.

5.2 From an agent view

Property 1 and 2 characterize overgeneralized and overspecific words by describing how their teaching influences the integration measures. However, this characterization can not be immediately used by an agent to find the right level of generality. Because one agent does not have access to the other agent's ontology, it can not compute how the teaching of a word influences the understandings rate. Therefore the agents follow the *expected increase in understandings rate*.

We use the notation $Exp(c, MPUR(m, \langle Ag_i, Ag_j\rangle))$ to refer to the expected value of $MPUR(m, \langle Ag_i, Ag_j\rangle)$, after the concept c has been taught. Given that the current $MPUR(m, \langle Ag_i, Ag_j\rangle)$ is 0, the expected value after c is taught can be calculated as follows (M_i is the meaning space in Ag_i's ontology, and M_j the meaning space in Ag_j's ontology)

- if $\mathcal{I}(c) = m$ then $Exp(c, MPUR(m, \langle Ag_i, Ag_j\rangle)) = 1$
- if m is a descendant of $\mathcal{I}(c)$ in M_i then $Exp(c, MPUR(m, \langle Ag_i, Ag_j\rangle))$ =Pr($\mathcal{I}(c)$ is not internal in M_j)
- if the first two conditions do not hold then $Exp(c, MPUR(m, \langle Ag_i, Ag_j\rangle)) = 0$

The first condition states that if c exactly means m, then the agent is certain that teaching the word c enables communication of the meaning m. The second condition states that, if c means something more general than m, the expected *MPUR* equals the probability that the other agent does not consider the word c overgeneralized. In our case this boils down to the probability that the meaning of c is a leaf in M_i, i.e. the ontology of Ag_j does not contain more specific concepts than c. The last condition states that, if c is not equal or more general than m, c can not be used to communicate m, and therefore the teaching of c will not increase the *MPUR* w.r.t. m.

The expected *PUR* (corresponding to definition 4) after c is taught can be calculated by averaging over the expected *MPUR*'s:

- Exp(c, $PUR(\langle Ag_i, Ag_j\rangle)$) = $\frac{1}{\#V_i} \sum_{m \in V_i} Exp(c, MPUR(m, \langle Ag_i, Ag_j\rangle))$

Because the agents must base their decision which meaning to select on *expectations*, the agents can not be certain that they find the right level of generality. Therefore, they must decide whether to attach more value to expected *MPUR*, or to expected *PUR*. This decision is set down in the parameters θ_1 and θ_2 which indicate the importance of respectively *MPUR*, and *PUR*. Using these parameters, the meaning that the meaning selection strategy selects is given by:

DEFINITION 6. *Given that Ag_i intends to communicate a meaning m. The meaning selection strategy is described by:* $argmax_{c \in \mathcal{C}_i}(\theta_1 \cdot Exp(c, MPUR(m, \langle Ag_i, Ag_j\rangle)) + \theta_2 \cdot Exp(c, PUR(\langle Ag_i, Ag_j\rangle)))$*, where:*

- θ_1 *is the importance factor for MPUR*
- θ_2 *is the importance factor for PUR*

In the next section we will investigate the effects of different importance factors for *MPUR* and *PUR*.

6. EXPERIMENTS

For our experiments, we adopt a group of 15 agents. An agent's ontology is randomly created according to $B_g = (3, 3, 3, 3, 3, 0)$ and $B = (2, 2, 2, 2, 1, 0)$, and contains 46 concepts. An experiment consists of t steps, where at each step a random speaker and hearer is selected from the group of agents, and a random concept from the speaker's ontology. We have prevented the same hearer-speaker-concept pair to be selected twice in the same experiment. The speaker communicates the concept to the hearer using a dialogue that conforms to the ANEMONE communication protocol (Figure 2) and a word selection strategy that selects the most frequently used word [5]. The speaker follows a meaning selection strategy that conforms to definition 6. After each step, we measure the following:

1. UR: the understandings rate, calculated according to definition 5.
2. Avg. Dialogue length : The average length of a dialogue of a randomly selected speaker-hearer-concept.
3. Avg. Nr. CDP : The average number of times that a concept is taught in (only) the CDP layer, in a dialogue of a randomly selected speaker-hearer-concept.
4. Avg. Nr. CEP : The average number of times that a concept is taught in the CEP layer, in a dialogue of a randomly selected speaker-hearer-concept.

In the next sections we will describe the results of six different experiments that were performed using different meaning selection strategies. To obtain statistical significance, we have performed every experiment 10 times of which we will present the mean outcomes. For all results, the standard deviation was less than 5 percent of the mean.

6.1 Agents that know the ontology model

In the previous section, we have argued that the speaker can determine the expected MPUR after teaching a concept by using the probability that the hearer's ontology contains no subconcepts of that concept. In this section, we assume that the agents know the ontology model, i.e. they know that $B = (2, 2, 2, 2, 1, 0)$ and $B_g = (3, 3, 3, 3, 3, 0)$. With this knowledge, an agent Ag_i can compute the probability that a meaning m is considered (non-) overgeneralized by an agent Ag_j as follows:

- if $ln(m)$ < *the depth of M* then Pr(m is internal in M$_j$) $= \prod_{i=0}^{ln(m)} \frac{b_i}{b_i^g}$
- if $ln(m)$ = *the depth of M* then Pr(m is internal in M$_j$) $= 0$
- Pr(m is not internal in M$_j$) = 1 - Pr(m is internal in M$_j$)

In these formulae, $b_0, .., b_d$ are typical elements of vector B, and $b_0^g, .., b_d^g$ are typical elements of B_g.

For example, in our experiments, the probability that a meaning at layer number 0 is internal is $\frac{2}{3}$. The probability that a meaning at layer number 4 is internal is $\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3}$. The probability that a meaning at layer number 5 is internal is 0.

A common pattern of dialogues in ANEMONE is that the speaker speaks a relatively general concept c, after which the hearer requests for specification, after which the speaker applies its meaning selection strategy a second time and speaks a more specific concept d. When the speaker applies the meaning selection strategy for the second time, it can use extra knowledge to compute the probability that d is considered overgeneralized by the hearer, namely that concept c *is* considered overgeneralized. We incorporate this idea in the meaning selection strategy using a conditional probability. An agent Ag_i that knows that a meaning n is overgeneralized for the hearer Ag_j computes the probability that a (more specific) meaning m is considered overgeneralized as follows:

$Pr(e1|e2) = \frac{Pr(e1)}{Pr(e2)}$, where

- $e1$ is the event that m is internal in M$_j$
- $e2$ is the event that n is internal in M$_j$, where n is an ancestor of m.

This can be proven as follows. According to Bayes theorem [14], $Pr(e1|e2) = \frac{Pr(e2|e1) \cdot Pr(e1)}{Pr(e2)}$. Note that $Pr(e2|e1)$ is 1, because $e2$ is implied by $e1$. Hence, $Pr(e1|e2) = \frac{Pr(e1)}{Pr(e2)}$.

Experiment 1

In the first experiment, we used parameters $\theta_1 = 1$ and $\theta_2 = 0$. In other words the agents only take the expected *MPUR* into account in their meaning selection strategy. Because they are only interested in the expected increase in MPUR concerning the meaning that they *currently* want to convey, we call this strategy a *short term strategy* (STS). The results of applying a short term strategy for 10000 steps is shown in Figure 3. The situations at 0 steps can be

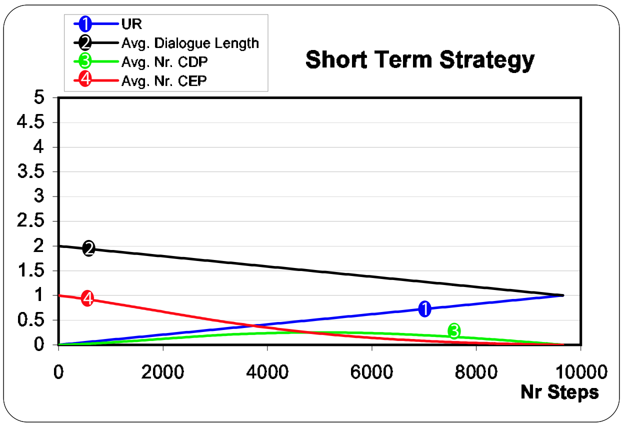

Figure 3: Results experiment 1

explained as follows. Because the agents have not taught any concepts to each other, no agent understands any other agent, hence UR is 0. This means that in every dialogue, the agents have to visit the CDP or CEP layer of the protocol. Because the agents do not share any words that they can use for giving concept definitions, all teaching of new words is done using CEP (where the meaning of a word is conveyed by pointing to shared instances). Hence Avg.Nr.CEP is 1 and Avg.Nr.CDP is 0. Because the agents visit the CEP layer every dialogue, the average dialogue length is 2.

As the number of steps increase, the agents teach concepts to each other, and the UR slowly increases. Also, the Avg.Nr.CDP increases because giving definitions becomes a viable option to teach new concepts, once a substantial amount of concepts is shared. As a result of this, there is less need for CEP, and the Avg. Nr CEP slowly decreases. Hence, the Avg. dialogue length also decreases.

Experiment 2

In experiment 2, we used parameters $\theta_1 = 0$ and $\theta_2 = 1$. In other words, the agents only take the expected *PUR* into account in their meaning selection strategy. Because they are interested in the expected increase in MPUR concerning any concept in their ontology, regardless whether they currently intend to convey it or not, we call this a *long term strategy* (LTS). The results of applying the long term strategy for 10000 steps is shown in Figure 4.

Using the long term strategy, the Avg.Nr.CEP is relatively high in the beginning. This is because the speaker may end up teaching three or four general concepts to the hearer, before it teaches the concept that is specific enough for the hearer to accept. As a result of this, the Avg. dialogue length is also relatively high. We can also observe that the strategy that aims at increasing the *PUR*, indeed gives rise to a fast increase in UR. Therefore, the Avg.Nr.CEP and Avg. Dialogue length decrease quickly in the beginning.

One of the reasons that experiment 2 exhibits a faster increase of UR than experiment 1 is that the Avg.Nr.CEP is higher in experiment 2 than in experiment 1. Another reason is that the concepts that are taught in experiment 1 are overspecific and therefore only increase UR a little (property

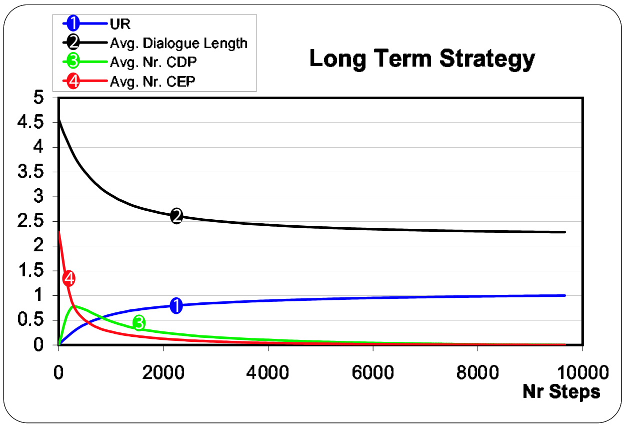

Figure 4: Results experiment 2

2). To support this claim we included Figure 6 where the strategies in experiments 1 and 2 (and 3) are compared in a graph with the total number of CEP on the x-axis. Furthermore, this figure reveals that the total number of CEP that is required to reach an UR of 1 is around 1300 using LTS, and around 5000 using STS. Therefore, the communication vocabulary that is produced by LTS is also much smaller than the CV that is produced by STS.

The following table compares the short term strategy (experiment 1) with the long term strategy (experiment 2).

	STS	LTS
Increase in UR	-	+
Initial Avg.Nr.CEP.	+	-
Avg. Dialogue Length	+	-

With respect to a fast increase in UR, the LTS performs better than the STS. However, the dialogues in the LTS are longer, and the Avg.Nr.CEP is high in the beginning. In the following experiment, we aim at achieving the best of both worlds.

Experiment 3

In experiment 3, we used parameters $\theta_1 = 1$ and $\theta_2 = 5$, such that the agents take the expected *MPUR* and *PUR* into account. Because it is a mixture of the short term strategy and the long term strategy, we call this the *medium term strategy* (MTS). The results are shown in Figure 5. As this

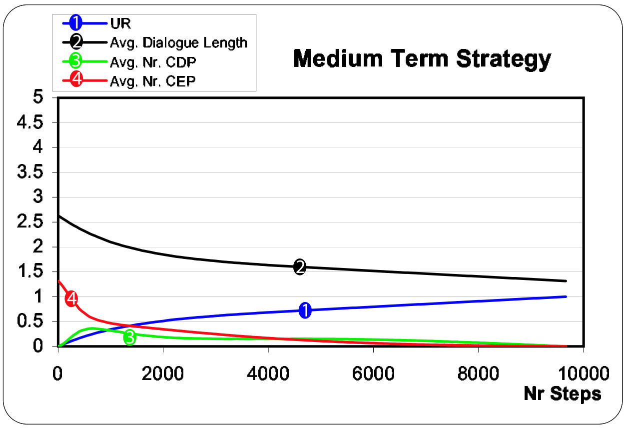

Figure 5: Results experiment 3

Figure reveals, the MTS gives rise to a faster increase of UR than the STS (experiment 1), and it gives rise to shorter dialogues and initial Avg.Nr.CEP than the LTS.

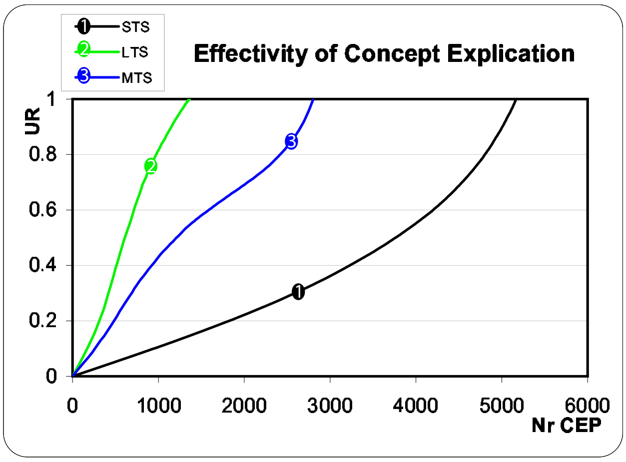

Figure 6: Comparison of experiments 1,2 and 3

6.2 Agents that learn the ontology model

The three experiments described in the previous section build on the assumption that the agents know the ontology model. In this section, we do not make this assumption, and make the agents learn the ontology model during their conversations. This is done as follows. For every meaning in its ontology, an agent keeps track of:

- N_1 the number of agents that regarded the meaning overgeneralized. These agents have responded "ReqSpec" to Inform-messages containing this meaning.
- N_2 the number of agents that did not regard the meaning overgeneralized. These agents have responded "OK" to inform-messages containing this meaning.

N_1 and N_2 are both initialized to 1. Using these values for meaning m, agent Ag_i can approximate the probability that m is internal in a meaning space M_j as follows:

- Pr(m is internal in M$_j$) $= \frac{N1}{N1+N2}$

Experiment 4,5,6

Experiments 4,5 and 6 were performed using STS, LTS and MTS respectively, with agents that learn the ontology model as they participate in conversations. Figure 7 shows the results of experiments 4,5 and 6 in a similar fashion as Figure 6. This figure reveals that STS in experiment 4 gives rise to very similar results as STS in experiment 1. This is because STS incites agents to select the most specific meaning. The inaccurate approximation of the ontology model in experiment 4, does change this strategy, as the agents will continue to select the most specific meaning anyway. The LTS incites the agents to select the most general meaning. Therefore, the LTS in experiment 5 gives rise to the same results as the LTS in experiment 2. The situation is different with the MTS, which incites agents to select a meaning that is a right balance between specificity and generality. An inaccurate approximation of the ontology model, does influence the results of the MTS, as can be seen when the results of experiment 3 are compared with experiment 6 in Figure 6 and 7.

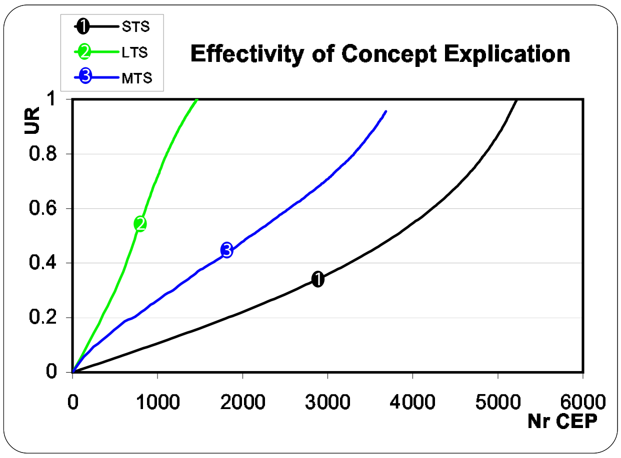

Figure 7: Comparison of experiments 4,5 and 6

7. CONCLUSION

In this paper, we have argued that finding the right level of generality is important for ontology negotiation. We have experimentally supported this claim by comparing different communication strategies that incite the agents to convey their information at different levels of generality. An agent that conveys information using a very specific word, runs the risk that the other agent does not know the word. An agent that conveys information using a very general word, runs the risk of being too vague which would result in a lengthy dialogue.

We have also shown that the agents can reliably assess the right level of generality themselves. They may do this by recording how many other agents do and do not consider a meaning overgeneralized. As an agent participates in conversations, it builds up a model of the other agents' ontologies, which enables it to find the right level of generality.

We believe that the communication strategies discussed in this paper are useful for agents in heterogeneous systems, as they prescribe which individual actions the agents must undertake in order to achieve the global goal of establishing an effective communication vocabulary. We intend to continue this line of research by incorporating tasks in the model. In such a model, the criteria of overgeneralization and overspecification become dependent on the tasks that the agents are discussing. Furthermore, we intend to enrich the ontologies of the agents with additional constructs such as attributes and part-of relations.

8. REFERENCES

[1] FIPA Ontology Service Specification. http://www.fipa.org/specs/fipa00086/.

[2] S. Bailin and W. Truszkowski. Ontology negotiation between intelligent information agents. *Knowledge Engineering Review*, 17(1):7–19, 2002.

[3] T. Bylander and B. Chandrasekaran. Generic tasks for knowledge-based reasoning: the right level of abstraction for knowledge acquisition. *Int. J. Man-Mach. Stud.*, 26(2):231–243, 1987.

[4] E. D. de Jong and L. Steels. A distributed learning algorithm for communication development. *Complex Systems*, 14(4), 2003.

[5] J. v. Diggelen, R. Beun, F. Dignum, R. v. Eijk, and J.-J. Meyer. A decentralized approach for establishing a shared communication vocabulary. *Proceedings of the International Workshop on Agent Mediated Knowledge Management (AMKM'05) held with AAMAS 2005.*

[6] J. v. Diggelen, R. Beun, F. Dignum, R. v. Eijk, and J.-J. Meyer. Optimal communication vocabularies and heterogeneous ontologies. In *Developments in Agent Communication*, LNAI 3396. Springer Verlag, 2004.

[7] J. v. Diggelen, R. Beun, F. Dignum, R. v. Eijk, and J.-J. Meyer. ANEMONE: An effective minimal ontology negotiation environment. *Proceedings of the Fifth International Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2006.

[8] S. A. Goldman and M. J. Kearns. On the complexity of teaching. *Journal of Computer and System Sciences*, 50(1):20–31, Feb. 1995.

[9] T. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.

[10] D. Malik and M. Sen. *Discrete Mathematical Structures: Theory and Applications*. Blackwell, 2004.

[11] E. Mena, A. Illarramendi, V. Kashyap, and A. Sheth. OBSERVER: An approach for query processing in global information systems based on interoperation across pre-existing ontologies. *International journal on Distributed And Parallel Databases (DAPD)*, 8(2):223–272, April 2000.

[12] N. F. Noy and M. A. Musen. Prompt: Algorithm and tool for automated ontology merging and alignment. *In Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 2000.

[13] A. D. Preece, K. ying Hui, W. A. Gray, P. Marti, T. J. M. Bench-Capon, D. M. Jones, and Z. Cui. The KRAFT architecture for knowledge fusion and transformation. *Knowledge Based Systems*, 13(2-3):113–120, 2000.

[14] J. Rice. *Mathematical Statictics and Data Analysis.* Duxbury Press, 1995.

[15] L. Steels. Language as a complex adaptive system. In M. Schoenauer, editor, *Proceedings of PPSN VI*, Lecture Notes in Computer Science. Springer-Verlag, Berlin, Germany, September 2000.

[16] M. Uschold and M. Gruninger. Creating semantically integrated communities on the world wide web. *Semantic Web Workshop Co-located with WWW 2002 Honolulu*, 2002.

[17] J. Wang and L. Gasser. Mutual online ontology alignment. *Proceedings of the Workshop on Ontologies in Agent Systems, held with AAMAS 2002.*

[18] A. Williams. Learning to share meaning in a multi-agent system. *Autonomous Agents and Multi-Agent Systems*, 8(2):165–193, 2004.