Ontological Feedback in Multiagent Systems

Robbert-Jan Beun, Rogier M. van Eijk and Huub Prüst
Information and Computing Sciences
Utrecht University, P.O. Box 80089
3508 TB Utrecht, The Netherlands
{rj, rogier, huub}@cs.uu.nl

Abstract

In this paper, we present a computational framework for the detection of ontological discrepancies in multiagent systems. The framework is developed as a basis for the generation of feedback utterances at the ontological level. In our method, presuppositions are extracted from the sender's message, expressed in a common vocabulary, and compared with the recipient's ontology, which is expressed in type theory. Discrepancies are detected by the receiving agent if it notices type conflicts, particular inconsistencies or ontological gaps. Depending on the kind of discrepancy, the agent generates a particular feedback message in order to establish alignment of its private ontology with the ontology of the sender.

keywords: ontological feedback, ontological discrepancies, presuppositions, type theory

1. Introduction

When agents transfer information, they need a conceptualisation of the domain of interest and a shared vocabulary to communicate facts with respect to this domain. The conceptualisation can be expressed in a so-called ontology, which is often defined in a formal language, such as a programming language or a formal logic. An ontology abstracts the essence of the domain of interest and helps to catalogue and distinguish various types of objects in the domain, their properties and relationships (see, e.g. [8, 13]).

In successful communication and in collaborative performance of tasks, agreement between different agents with respect to the ontology is crucial or, at least, the agents should be aware of existing discrepancies. In case of discrepancies, various troublesome situations may arise. One of the agents, for instance, by mistake may assume that a particular concept is shared, while the other has no knowledge about it. Or worse, both dialogue partners have different conceptualisations, while the relevant discrepancy remains unnoticed. Ontological discrepancies may cause serious communication flaws and the generation of adequate feedback in order to avoid these flaws is an essential part of modelling a proper communication process.

In principle, there is a range of approaches to achieve ontological agreement. On the one side, developers can agree in advance upon a standard domain ontology and embed it in all future ontology design [14]. On the other side, we may accept the existence of ontological variations and design a mechanism that solves discrepancies during the communication process (see e.g. [9, 19]). Since multiagent systems are often developed by multiple parties and can be considered as open and distributed systems, we believe (in line with [23]) that it is hard to imagine that there will ever be a consensus about ontologies for every possible domain of interest. Therefore, in the present paper we will adopt the second approach and concentrate on communication protocols to solve ontological discrepancies between agents.

The aim of this paper is to develop a computational framework for the detection of ontological discrepancies in multiagent systems. In particular, it will be shown how ontological discrepancies can be detected during a communicative situation and how a dialogue participant can react to these observed discrepancies. Agents may detect discrepancies by, for instance, type conflicts, ontological gaps and particular inconsistencies that emerge during the conversational process. Depending on the kind of discrepancy, the agent generates a particular feedback message in order to establish alignment of its private ontology with the ontology of the sender. For that, we will adopt an approach in which agents have a dynamic mental state that contains ontological information about the domain of interest in terms of simple type theoretical contexts.

In Section 2, we will consider feedback in communication. In Section 3, we will discuss ontological discrepancies. In Section 4, we will turn to the basic notions of the communicative situation and in Section 5, we will introduce the fragment of type theory that is used to express the ontological state of an agent. In Section 6, the actual mechanism is described and some examples are shown to demonstrate the feasibility of the system. We wrap up in Section 7 and provide directions for further research.

2. Feedback in communication

In communication, feedback is used for a broad range of responses at various levels and has an enormous diversity, varying from a simple nod in human-human communication or a particular bit that indicates the receipt of a message in computer-computer communication to a written comment that evaluates the quality of a scientific paper. However, for various reasons, we have no accurate mathematical theory for adequate communicative behavior and the application of cybernetic models to communicative activities has only a limited scope of relevance [22].

When we look at general feedback phenomena in conversations between humans, sequences in terms of speech acts appear to be rather chaotic and seem hardly subjected to any rules. Questions can be followed by answers, denials of the relevance of the question, rejections of the presuppositions of the question, statements of ignorance, and so on (see e.g. [17]). An example of general rules for cooperative contributions, and conversational feedback in particular, are the Gricean maxims for conversation, such as 'tell the truth' (quality), 'say enough, but not too much' (quantity), 'be relevant' (relevance) and 'use the appropriate form' (manner) ([12]).

In the field of multiagent systems, there are no general theories specifying what the most adequate feedback sequences are in a given communication setting. Most frameworks allow the possibility of giving feedback, but do not specify which feedback and how this should be done. The agent communication languages KQML [10] and FIPA ACL [11] include performatives for giving feedback. For instance, in KQML, there are feedback performatives reporting malformed messages ('error'-performative) or an inability to respond ('sorry'-performative). Furthermore, FIPA ACL includes the performative 'not understood' to report that a previous message has not been understood. The description of this performative states that:

> ... the content part of the message contains the reason for the failure to understand. For instance, an agent did not understand a message because it did not recognize the ontology. There is no guarantee that the reason is represented in a way that the receiving agent will understand. However, a co-operative agent will attempt to explain the misunderstanding constructively [11].

In other words, FIPA ACL allows the possibility for giving feedback, but the actual generation mechanism remains unspecified. This implies that the generation of adequate feedback sequences to a certain extent is left to the responsibility of the system developer.

Suppose, for instance, that we have two agents A and B that have two separate ontologies and that A asks the question: 'Is this vessel alive?'. B's ontology contains, among other things, a representation for the words 'vessel' and 'alive' (and knows which vessel is meant), but believes that vessels are inanimate and that alive can only be applied to the category of animate objects (including whales). Assuming that the agent should be relevant and truthful, then what should the response of agent B be? Clearly, we have abundant possibilities for feedback:

$B1$: What do you mean by 'alive'?
$B2$: What do you mean by 'vessel'?
$B3$: 'Alive' is only applicable to animate things
$B4$: 'Alive' is not applicable to vessels
$B5$: 'Alive' is not applicable to inanimate things
$B6$: Vessels are not whales
$B7$: Inanimate things are not animate things
$B8$: Vessels are not animate things
$B9$: Can vessels be alive?
$B10$: Are vessels whales?
and so on ...

Which utterance is the most adequate one and which rules we should apply in order to generate these feedback sequences depends among others on what the agents know of each other like the beliefs they already share with each other. In this paper, feedback will be based on the type of discrepancy and we will abstract from the various types of beliefs and roles of the agents.

3. Ontological discrepancies

In order to provide feedback at the ontological level, we need a computational decision criterium that tells us whether two ontologies are incompatible. This is not a trivial matter, however. We will first have a brief informal discussion that motivates the rules presented in Section 6.

In many frameworks (like [16, 25]), discrepancies between two ontologies are *implicitly* defined relative to a third ontology, namely the ontology of an external observer. In [25], Visser *et al.* give an overview of different ontological mismatches, such as 'structure mismatch', 'attribute-assignment mismatch' and 'concept-and-term mismatch'. The 'concept-and-term mismatch', for instance, occurs when two ontologies use the same definiens but differ in both the concept they define and the term linked to the definiens.

To explain this notion, in [25], the following ontologies

O_1 and O_2 are given (we adopt the same notation):

$$O_1 : vessel(X) \leftarrow seagoing(X) \wedge large(X)$$

$$O_2 : whale(X) \leftarrow seagoing(X) \wedge large(X)$$

According to Visser *et al.* there is a 'concept-and-term mismatch' between O_1 and O_2: the ontologies use the same definiens to define different concepts. But is this really so? Can we also conclude that they define different concepts? Here, it is important to observe that in the answer of Visser *et al.* implicit reference is made to a third ontology. It is assumed that external observers have an ontology of their own about whales and vessels and know what their similarities and differences are. For instance, in our own ontology vessels are characterised to be inanimate objects while whales are animate objects. So, from our 'god's eye' point of view, there is a discrepancy: whales and vessels constitute different concepts in our ontology. In other words, the decision criterium comes from information of the third observer.

In multiagent systems, however, the assumption of a 'god's eye' view is not realistic in many cases. Agents are distributed entities that autonomously operate in their environment, of which they have their own subjective view. As humans, they need to detect ontological discrepancies by *themselves*, on the basis of their own subjective view on the world and not on the basis of information hidden in a third observer.

Now, assume that O_1 and O_2 only contain the information above. Can we conclude that they define different concepts? From the local perspective of an agent, it cannot be concluded that there is a discrepancy. Both ontologies account for large seagoing objects, and indeed, there are tasks in which the distinction between vessels and whales is of no importance; e.g., in navigating a ship where large seagoing objects have to be avoided. The only conflict here seems to be a *lexical* discrepancy.

Consider the refined ontologies O_3 and O_4 that both define the categories of animate an inanimate objects, which are disjoint:

$$O_3 : animate(X) \leftrightarrow \neg inanimate(X)$$
$$vessel(X) \leftarrow seagoing(X) \wedge large(X) \wedge inanimate(X)$$

$$O_4 : animate(X) \leftrightarrow \neg inanimate(X)$$
$$vessel(X) \leftarrow seagoing(X) \wedge large(X) \wedge animate(X)$$

From the assumption that both agents believe that objects cannot be both animate and inanimate, we conclude that there is an ontological discrepancy between O_3 and O_4. Note that no reference to a third ontology needs to be made to characterise this discrepancy.

Below, the decision criteria for discrepancies will be expressed in terms of type theory where the addition of particular information to ontologies yields so called legal or illegal contexts. A legal context is a context where the addition

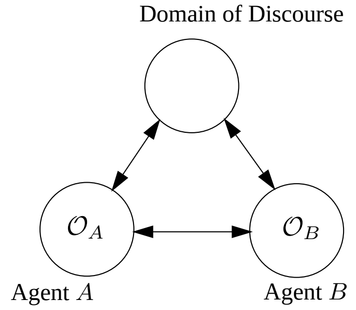

Figure 1. The triangle metaphor

of new information was adhered to the rules of the type system (see also Section 5). For instance, the introduction of new predicates is only possible if the type of its arguments is already included in the ontology; otherwise, the context is illegal. If we would like to add, for instance, the assumption that 'whales are heavy', we must at least have a notion of 'heaviness' and 'whale' in our current state.

In the detection stage of the interpretation process of an incoming message, particular information – so called presuppositions – is extracted from the message. In cases where the addition of presuppositions to the ontology of the receiver yields an illegal context, the receiving agent has to generate particular feedback.

4. Dialogue model

The dialogue framework employed in this paper is based on a simple model of interaction [2, 21]. Underlying this model is the recognition that agents interact naturally with their environment in two ways: symbolically and physically. On the one hand, if there is an intermediary interpreter, agents can interact symbolically and use a communication language to give commands, ask for or provide information, etcetera. On the other hand, physically, one manipulates objects, for instance, by moving or fastening them, or observes them by seeing, feeling, hearing, tasting or smelling.

The distinctive interaction channels are represented in the so-called triangle metaphor (Figure 1), where the corners represent the domain of discourse (or the external world) and the two agents, and the arrows the flow of information between the agents themselves and between the agents and the domain.

In the model, we clearly distinguish between the world and knowledge (or beliefs) about the world. The world is represented by a set of concrete objects that are instances of particular classes (vessels, whales, persons, . . .) which have particular characteristics (colour, weight, . . .) with a particular value (green, red, heavy,. . .). Also, the objects may have particular relations between them (next, heavier, . . .). The knowledge about the world is an explicit representation of the objects with their characteristics and their relations. We will use the term *ontology* (indicated by $\mathcal{O}_A$ and $\mathcal{O}_B$ in the figure) to indicate the abstraction by the agent of the relevant aspects of the world that help to classify and distinguish various types of objects in the domain, their properties and relationships in order to express particular facts about the world. The ontology expresses the possible and necessary characteristics of the domain. To indicate the subjective nature of the agents' information state, we will often use the term 'belief' instead of knowledge.

In the dialogue process, we will make the following assumptions with respect to the dialogue model. Firstly, we assume that both agents have access to the objects in the domain of discourse, i.e. they can observe pointing actions from their dialogue partner and they are able to link these pointing acts to mental objects. This does not mean, however, that all features of the object are perceivable. Secondly, we will assume that if one of the agents points to an object, the other observes the same object.

5. Ontologies in type theory

In this paper, the ontology will be expressed in type theory (TT). TT, which is actually based on typed lambda calculus, is a powerful logical formalism in the field of theorem proving and programming languages. In the field of agent communication, TT was used in the DenK-project [2, 5] as a knowledge representation to model various types of beliefs. In the project, an 'intelligent' agent was modelled that supported a human user in its use of a particular domain. Although the system was applied to the domain of an electron microscope, it was intended to be generic in that its architecture and various techniques for knowledge representation and construction were independent of the field of application. The agent's belief states, such as private and common beliefs, were modelled as type theoretical contexts. In the DenK-project the formalism was used for modelling the ontological assumptions and beliefs about the task domain (the electron microscope) and the cognitive dynamics of the agent's belief state, in particular, the change of beliefs as a result of domain observations and dialogue contributions by the user.

Apart from its intrinsic dynamic properties, TT has important advantages over formalisms such as predicate logic or discourse representation theory. First, and at the heart of type theory, is the formal distinction between objects and types. Types often represent a particular concept, whereas objects can be considered as instances of these concepts. In type theory it can be expressed, for instance, that we have the concept 'whale' and that 'willy' is an instance of the concept whale. Second, TT embodies the notion of contextuality (or sequentiality) that tells us that new assumptions can only be added as long as the current belief state satisfies particular constraints. To some extent, the notion of sequentiality corresponds to the notion of presupposition used in linguistics [24], where, for instance, the sentence 'Type theory is not the most prominent logical framework that agent researchers use when they model ontologies', presupposes among other things 'the existence of a particular instance of the most prominent logical framework that ... ' (namely description logic), 'the existence of agent researchers and that they may use or model something', 'the existence of ontologies and that they can be modelled'. An important characteristic of presuppositions is that they can be inferred from both the sentence and its interrogative form.

Here we will only give an informal introduction and show how particular ontological information can be expressed in a very limited fragment of TT (see e.g. [1, 18] for comprehensive introductions). Beliefs of an agent can be represented in TT as so called contexts; contexts consist of sequences of expressions and list everything that has been assumed so far and everything that has explicitly been inferred from these assumptions. The building blocks in TT are expressions of the form

$$G : H \tag{1}$$

which indicate that an object G has type H. These expressions are called statements. We can express, for instance, that a particular object (namely *willy*) is of type *whale*:

$$willy : whale \tag{2}$$

The object *willy* is called an 'inhabitant' of the type *whale*. Concepts, such as whales, are expressed as a special type named 'sort' and are represented by '$\star_s$' in (3):

$$whale : \star_s \tag{3}$$

Variables like *willy* and *whale* are referred to as 'type variables'.

The notion of sequentiality plays a role in the order in which statements can be added. For instance, (2) cannot be added to the context, unless the belief state of the agent already contains statement (3).

Predicates are considered as a functions that take a particular type as an argument and that yield a truth value. For that reason, we also have to introduce propositions; we will do this by introducing a new kind of sort '$\star_p$' For instance:

$$animate : \star_s \tag{4}$$
$$alive : animate \to \star_p \tag{5}$$

expresses that the predicate alive takes an animate object as an argument (note that we first explicitly have to introduce the existence of the concept animate).

Inheritance is introduced by the subsumption operator '$<$', which indicates that inhabitants of a more specific type can be applied in every case where inhabitants of the more general type may be applied (where types on the left are lower in the hierarchy):

$$\mathit{whale} < \mathit{animate} : \star_s \quad (6)$$
$$\mathit{vessel} < \mathit{inanimate} : \star_s \quad (7)$$

Similarly, types $\star_s$, $\star_p$ and $A \rightarrow \star_p$ themselves are inhabitants of the top-level type $\Box$, so for instance:

$$\star_s : \Box \quad (8)$$

Contexts are open to new information and can be extended as long as new introductions are adhered to the rules of the type system; these contexts are called *legal contexts*. So, given a particular context, the rules of the type system constrain the way in which statements can be combined into new legal statements. This can be expressed by so called judgements:

$$\Gamma \vdash E : T \quad (9)$$

which expresses that term E has type T, given the assumptions in context Γ.

A context Γ is defined to be *legal* if $\Gamma \vdash \star_s : \Box$.

The rules of the type system are as follows (cf. [1, 15]).

$$\frac{}{\epsilon \vdash \star : \Box}(\mathit{axiom}) \quad (10)$$

where $\star$ can stand for $\star_s$ or $\star_p$. This axiom defines the empty context (denoted by ϵ) to be legal.

$$\frac{\Gamma \vdash A : s}{\Gamma, x : A \vdash x : A}(\mathit{start}) \quad (11)$$

$$\frac{\Gamma \vdash y : \star_s}{\Gamma, x < y : \star_s \vdash x : \star_s}(\mathit{start}') \quad (12)$$

$$\frac{\Gamma \vdash y : \star_s}{\Gamma, x < y : \star_s \vdash x < y : \star_s}(\mathit{start}'') \quad (13)$$

where x is fresh in Γ and s can stand for $\star_s$, $\star_p$ or $\Box$. These rules allow us to extend a context with a statement.

$$\frac{\Gamma \vdash A : B \quad \Gamma \vdash C : s}{\Gamma, x : C \vdash A : B}(\mathit{weakening}) \quad (14)$$

$$\frac{\Gamma \vdash A : B \quad \Gamma \vdash y : \star_s}{\Gamma, x < y : \star_s \vdash A : B}(\mathit{weakening}') \quad (15)$$

where x is fresh in Γ. These rules guarantee that context extensions are monotonic in the sense that statements remain derivable after an extension of the context.

$$\frac{\Gamma \vdash A : \star_s}{\Gamma \vdash A \rightarrow \star_p : \Box}(\mathit{predicate}) \quad (16)$$

This rule defines the formation of predicate types.

$$\frac{\Gamma \vdash y : A \quad \Gamma \vdash A < B : \star_s}{\Gamma \vdash y : B}(\mathit{subsume}) \quad (17)$$

$$\frac{\Gamma \vdash C : B \rightarrow \star_p \quad \Gamma \vdash A < B : \star_s}{\Gamma \vdash C : A \rightarrow \star_p}(\mathit{subsume}') \quad (18)$$

These rules specify that variables of a particular subtype also belong to the supertype, and that predicates applicable to the supertype are also applicable to a subtype.

$$\frac{\Gamma \vdash F : A \rightarrow \star_p \quad \Gamma \vdash a : A}{\Gamma \vdash F(a) : \star_p}(\mathit{apply}) \quad (19)$$

This rule states that the application of a predicate to an instance of the argument type is of type 'proposition'.

Hence, according to these rules the context Γ consisting of (4), (5), (6), (2) can be shown to be legal. Moreover, using rule (17), Γ can be legally extended with:

$$\mathit{willy} : \mathit{animate} \quad (20)$$

and according to rule (19) the resulting context can be legally extended with:

$$\mathit{alive}(\mathit{willy}) : \star_p \quad (21)$$

Note that the context only states that *alive*(*willy*) is a proposition, not that the proposition is true. In type theory, propositions are considered as types themselves and a proposition is true if there exists an object with the proposition as its type (e.g., $p_{17} : \mathit{alive}(\mathit{willy})$). However, since truth values of propositions are not essential for our treatment of ontological discrepancies we will not further consider this issue.

Below the notion of presuppositions will be used to compare the incoming information from the communication language with the ontology of the receiving agent. Presuppositions will be expressed as lists of TT-statements. We will say that a list of presuppositions is *legal with respect to a particular context* if this context extended with the presupposition list is still a legal context, otherwise we speak of an illegal list.

6. Rules for detection and feedback

We define a FIPA-compliant agent communication language. We distinguish between messages to ask and answer questions about the state of the domain and messages for giving feedback at the ontological level.

Definition 1 We assume a vocabulary or lexicon of words, with typical elements w_1 and w_2. We define the following messages concerning the state of the domain:

$$\langle A, \text{QUERY-IF}(B, \varphi)\rangle$$
$$\langle A, \text{CONFIRM}(B, \varphi)\rangle$$
$$\langle A, \text{DISCONFIRM}(B, \varphi)\rangle$$

where A denotes the sender and B the recipient and φ is of the form 'this w_1 is a w_2' or 'this w_1 is w_2'.

The messages are used to query, confirm and disconfirm that φ is believed to be true of the domain, respectively. The content formula 'this w_1 is a w_2' expresses that the particular object under discussion (denoted by 'this w_1') is an instance of the category w_2. The content formula 'this w_1 is w_2' expresses that the particular object under discussion (denoted by 'this w_1') satisfies the predicate w_2.

Note the distinction between the semantic and pragmatic aspects of the communication language. The semantics deals with how words (w_1 and w_2) are interpreted in terms of the underlying ontology. The pragmatics deals with the aspects that are independent of the ontology; e.g., the message type (query, confirm and disconfirm) and the handling of the definite article 'this' and the indefinite 'a'.

Definition 2 We define the following feedback messages concerning the ontological level:

$$\langle A, \text{QUERY-IF}(B, \Gamma)\rangle$$
$$\langle A, \text{INFORM}(B, New_A(\Gamma))\rangle$$
$$\langle A, \text{INFORM}(B, Bel_A(\Gamma) \wedge Bel_B(\Gamma'))\rangle$$

where Γ and Γ' are type theoretical contexts.

The first of these messages is used to ask whether Γ is part of the recipient's ontology, the second is used to report that Γ is new to the sender (i.e., that it was not part of its ontology) and the third is used to indicate a mismatch between part Γ of the sender's ontology and part Γ' of the recipient's ontology.

We identify the following aspects of an agent.

Definition 3 An agent is a tuple (Γ, I) consisting of a type-theoretical context Γ defining the ontology, and an interpretation $I \subseteq Voc \times Var$ that maps words of the communication vocabulary (collected in the set Voc) to the type variables of the ontology (collected in the set Var). We use the notation $I(w) = \{t \mid (w, t) \in I\}$.

Note that words in the lexicon may be ambiguous (i.e. one word in the communication language is mapped to different type variables in the ontology) and words may have synonyms (i.e. different words in the communication language may be mapped to one term in the ontology).

Dialogues are started with a query of the form $\langle A, \text{QUERY-IF}(B, \varphi)\rangle$. We discern the following stages in the handling of the query by the recipient B.

(a) Derivation of ontological presuppositions

(b) Detection of ontological discrepancies

(c) Generation of a response

The first stage concerns the extraction of the ontological presuppositions that are part of the query. This is done on the basis of the receiving agent's interpretation and the pragmatics of the communication language. The result is a set of presupposition lists.

Definition 4 Given an interpretation I, the function F that generates a set of presupposition lists, is defined as follows:

$$F(\text{this } w_1 \text{ is a } w_2) \\ = \\ \{[C : \star_s, x_{17} : C, D < C : \star_s] \mid (C, D) \in I(w_1) \times I(w_2)\}$$

$$F(\text{this } w_1 \text{ is } w_2) \\ = \\ \{[C : \star_s, x_{17} : C, P : C \to \star_p] \mid (C, P) \in I(w_1) \times I(w_2)\}$$

Here x_{17} represents the particular object that is under discussion. (Recall our assumption that agents can observe pointing actions and can link them to mental objects.) Note that an ontological presupposition of the sentence 'this w_1 is a w_2' is that the category denoted by w_1 subsumes the category denoted by w_2. A presupposition of 'this w_1 is w_2' is that the predicate denoted by w_2 is applicable to the category denoted by w_1.[1]

From now on, we will assume that I is a one-to-one relation between Var en Voc, which implies that F yields only one presupposition list. In order to detect possible ontological discrepancies, the presupposition list is compared with the ontology of the receiver.

Definition 5 Given an ontology Γ (assumed to be legal) and a presupposition list X, we discern the following cases:

1. $\Gamma \vdash X$ (i.e., each statement of X is derivable from Γ)
2. $\Gamma \nvdash X$ and X is legal with respect to Γ
3. $\Gamma \nvdash X$ and X is illegal with respect to Γ, and there is a legal context Δ such that Δ, X is legal with respect to Γ
4. $\Gamma \nvdash X$ and X is illegal with respect to Γ and there does not exist a legal context Δ such that Δ, X is legal with respect to Γ.

In case (1), the presuppositions follow from the ontology; there is no discrepancy. In case (2), the presuppositions constitute a legal extension of the ontology. In case (3), there exists a gap which is bridged by the type theoretical context Δ (cf. [20]). Also, in case (4), there is a gap, but there is no bridging context.

Let the ontology Γ be defined as:

$$animate : \star_s$$
$$alive : animate \to \star_p$$
$$x : animate$$
$$vessel : \star_s$$

1 Note that more than one list of presuppositions is obtained when (some of) the words in the question are ambiguous. Some of them may be in line with the agent's own ontology and others not. When the agent cannot decide which list is most appropriate it has to generate ambiguity feedback. A useful strategy could be, for instance, that the receiver - who now becomes the sender in the subsequent turn - first considers possible synonyms and returns them to the questioner ('Do you mean x or y?').

Let the interpretation I be simply defined as:

$$\{(\text{alive}, alive), (\text{vessel}, vessel), (\text{animate}, animate)\}$$

We consider the following four instantiations of $\langle A, \text{QUERY-IF}(B, \varphi_i)\rangle$:

$$\begin{aligned} \varphi_1 &= \text{this } animate \text{ is } alive \\ \varphi_2 &= \text{this } animate \text{ is a } vessel \\ \varphi_3 &= \text{this } vessel \text{ is } alive \\ \varphi_4 &= \text{this } vessel \text{ is } animate \end{aligned}$$

- First of all, $F(\varphi_1)$ is given by:

$$\begin{aligned} &animate : \star_s \\ &alive : animate \to \star_p \\ &x : animate \end{aligned}$$

  Since $\Gamma \vdash F(\varphi_1)$ there is no discrepancy.

- Secondly, $F(\varphi_2)$ is given by:

$$\begin{aligned} &animate : \star_s \\ &x : animate \\ &vessel < animate : \star_s \end{aligned}$$

  In this case $\Gamma \nvdash vessel < animate : \star_s$ and thus $\Gamma \nvdash F(\varphi_2)$. However, $F(\varphi_2)$ is legal with respect to Γ.

- Thirdly, $F(\varphi_3)$ is given by:

$$\begin{aligned} &vessel : \star_s \\ &alive : vessel \to \star_p \\ &x : vessel \end{aligned}$$

  In this case $\Gamma \nvdash x : vessel$ and $\Gamma \nvdash alive : vessel \to \star_p$. According to rules (17) and (18) of the type system it is possible for an object x to be of two different types (i.e., *vessel* and *animate*) and for a predicate *alive* to be applicable to both *vessel* and *animate*, if one of these types subsumes the other. The expression $\Delta = vessel < animate : \star_s$ bridges the gap between Γ and X; that is, $\Delta, F(\varphi_3)$ is legal with respect to Γ. Note that this bridge is not unique; e.g., $animate < vessel : \star_s$ also bridges this gap.

- Finally, $F(\varphi_4)$ is given by:

$$\begin{aligned} &vessel : \star_s \\ &animate : vessel \to \star_p \\ &x : vessel \end{aligned}$$

  Since, concepts and predicates are incompatible types, which means that *animate* cannot be both of type '$\star_s$' and of type '$vessel \to \star_p$', we conclude that $\Delta, F(\varphi_4)$ is illegal with respect to Γ for any Δ.

What ontological feedback should be generated in each of the four cases? In case φ_1, there are no ontological discrepancies. The rule is to answer the question (for instance, in accordance with the Gricean maxims of cooperation). So, a response could be:

$$\langle B, \text{CONFIRM}(A, (\text{this } animate \text{ is } alive))\rangle$$

In case φ_2, the presuppositions are not part of the current context, but can be added to yield a legal context. Here the rule is to report that the newly derived ontological facts were unknown. So, the response could be to state that the ontological fact that the category of animates subsumes the category of vessels is new to agent B:

$$\langle B, \text{INFORM}(A, New_B(vessel < animate : \star_s))\rangle$$

In case φ_3, the presuppositions are not part of the agent's ontology and cannot be added to yield a legal context. However, there exists a context that can bridge the gap between the ontology and the presuppositions of the question. So, the rule is to ask whether a particular bridge is part of A's ontology. So, the response could be to ask whether the category of animates subsumes the category of vessels,

$$\langle B, \text{QUERY-IF}(A, vessel < animate : \star_s))\rangle$$

In case φ_4, a conflict has been detected. The rule here is to state the reason for the discrepancy. So, it could be stated that according to A animate should be a predicate, but according to B it is a category.

$$\langle B, \text{INFORM}(A, Bel_A(animate : vessel \to \star_p) \wedge Bel_B(animate : \star_s))\rangle$$

On the basis of the given feedback, agents can subsequently engage themselves in a dialogue discussing their ontologies. The particular rules that govern the subsequent dialogue will be investigated in future research.

7. Discussion

Essential to the approach presented in this paper is that ontological discrepancies are treated at the level of agents themselves, without the aid of an external observer. The framework accounts for the detection and handling of ontological discrepancies by the agents themselves, on the basis of their own subjective view on the world. This means that there is no reference to any (implicit) third ontology. It also means that the framework abstracts from a notion of truth which is inherent to model-theoretic approaches. Agents work towards agreement on the basis of their belief states and communicative acts. We believe that this approach is both theoretically and practically important for multiagent systems. The account has been worked out for a core ontology on the basis of simple type theoretical fragments.

Future work will include the extension of the basic framework to richer ontologies and more complex type theoretical constructs. Important research questions are (in line with [3, 4]): How does the initial feedback change the belief state of the agents and and how can subsequent dialogue contributions be modelled? To answer these questions we must consider a variety of additional constructs such as agent roles and various types of beliefs and speech acts. Furthermore, we believe that our approach is applicable to important other aspects of communication in multi-agent systems, such as argumentation and negotiation. We also expect the approach to be fruitful with regard to ontology alignment and merging [6, 26].

Acknowledgements

The authors would like to thank Dave Clark and the anonymous referees for their valuable comments on an earlier draft of this paper.

References

[1] R. Ahn. *Agents, Objects and Events: A Computational Approach to Knowledge, Observation and Communication*. PhD thesis, Technische Universiteit Eindhoven, the Netherlands, 2001.

[2] R. Ahn, R. Beun, T. Borghuis, H. Bunt, and C. v. Overveld. The DenK-architecture: A fundamental approach to user-interfaces. *Artificial Intelligence Review*, 8:431–445, 1995.

[3] R. Beun. On the generation of coherent dialogue: A computational approach. *Pragmatics and Cognition*, 9(1):37–68, 2001.

[4] R. Beun and R. v. Eijk. A cooperative dialogue game for resolving ontological discrepancies. In F. Dignum, editor, *Advances in Agent Communication*, volume 2922 of *Lecture Notes in Artificial Intelligence*, 2004.

[5] H. Bunt, R. Ahn, R. Beun, T. Borghuis, and K. v. Overveld. Multimodal cooperation with the DenK system. In H. Bunt, R. Beun, and T. Borghuis, editors, *Multimodal Human-Computer Communication*, volume 1374 of *Lecture Notes in Computer Science*, pages 39–67. Springer-Verlag, 1998.

[6] M. Burnstein, D. McDermott, D. Smith, and S. Westfold. Derivation of glue code for agent interoperation. *Autonomous Agents and Multi-Agent Systems*, 6(3):265–286, 2003.

[7] C. Castelfranchi and W. Lewis Johnson, editors. *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems*, New York, New York, 2002. ACM Press.

[8] B. Chandrasekaran, J. Josephson, and V. Benjamins. What are ontologies and why do we need them? *IEEE Intelligent Systems*, 14(1):20–26, 1999.

[9] R. v. Eijk, F. d. Boer, W. v. d. Hoek, and J.-J. Meyer. On dynamically generated ontology translators in agent communication. *International Journal of Intelligent Systems*, 16(5):587–607, 2001.

[10] T. Finin, D. McKay, R. Fritzson, and R. McEntire. KQML: An Information and Knowledge Exchange Protocol. In K. Fuchi and T. Yokoi, editors, *Knowledge Building and Knowledge Sharing*. Ohmsha and IOS Press, 1994.

[11] FIPA. Foundation for intelligent physical agents. Communicative act library specification. http://www.fipa.org, 2002.

[12] H. Grice. Logic and conversation. In P. Cole and J. Morgan, editors, *Speech Acts. Syntax and Semantics, Vol. 11*, pages 41–58. Academic Press, New York, 1975.

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

[14] C. Holsapple and K. Joshi. A collaborative approach to ontology design. *Communications of the ACM*, 45(2):42–47, 2002.

[15] L. Kievit. *Context-Driven Natural Language Interpretation*. PhD thesis, Katholieke Universiteit Brabant, the Netherlands, 1998.

[16] W. Kim and J. Seo. Classifying schematic and data heterogeneity in multidatabase systems. *Computer*, 24(12):12–18, 1991.

[17] S. Levinson. *Pragmatics*. Cambridge University Press, Cambridge, 1983.

[18] Z. Luo. *Computation and Reasoning: A Type Theory for Computer Science*. International Series of Monographs on Computer Science. Clarendon Press, Oxford, 1994.

[19] M. Obitko and V. Marik. Mapping between ontologies in agent communication. In V. Marik, J. Müller, and M. Pechoucek, editors, *Proceedings of CEEMAS 2003*, volume 2691 of *LNAI*, pages 191–203. Springer-Verlag, 2003.

[20] P. Piwek and E. Krahmer. Presuppositions in context: Constructing bridges. In P. Bonzon, M. Cavalcanti, and R. Nossum, editors, *Formal Aspects of Context*, volume 20 of *Applied Logic Series*. Kluwer Academic Publishers, 2000.

[21] C. Rich, N. Lesh, J. Rickel, and G. Garland. A plug-in architecture for generating collaborative agent responses. In Castelfranchi and Lewis Johnson [7], pages 782–789.

[22] A. Spink and T. Saracevic. Human-computer interaction in information retrieval: Nature and manifestation of feedback. *Interacting with Computers*, 10:249–267, 1998.

[23] L. Steels. The origins of ontologies and communication conventions in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 1(2):169–194, 1998.

[24] P. Strawson. On referring. *Mind*, 59:320–344, 1950.

[25] P. Visser, D. Jones, T. Bench-Capon, and M. Shave. Assessing heterogeneity by classifying ontology mismatches. In N. Guarino, editor, *Formal Ontology in Information Systems*, pages 148–162, Amsterdam, the Netherlands, 1998. IOS Press.

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