Argumentation and agreement over concept meaning in simple scenarios

Kemo Adrian
IIIA, Artificial Intelligence Research Institute CSIC,
Spanish Council for Scientific Research Campus UAB,
08193 Bellaterra, Catalonia (Spain)
Email: kemo.adrian@iiia.csic.es

***Abstract*—Disagreements over the meaning of symbols can occur in mutliagent systems when two agents have different classifications of their environment. We argue that a semiotic approach to multiagent communication protocols can help to overcome those disagreements. A triadic relation between a sign, its intentional and its extensional definition allows us to use argumentation theory in order to modify concepts dynamically. Following this idea, we propose a simple protocol that allows agents to solve disagreements over meaning in basic scenarios. Finally, we describe three of those scenarios and discuss their limitations and possible enhancements.**

I. Introduction

Our computers, as Universal Turing Machines, are designed to process symbols. However, when it comes to interface them with reality – or at least our human reality – computer scientists have to face a major issue. This issue is the definition of symbols' meaning, the symbol grounding problem. This issue became popular with Searle's Chinese room experiment [1] and was one of the major issues in Artificial Intelligence during the nineteen eighties. Today the question of symbol grounding is claimed to be settled: its achievement is possible. But with the ability to ground symbols arose a new problem in Artificial Intelligence. How can we allow different intelligent systems to communicate if they have grounded different symbol-meaning associations? Since this issue has already been investigated by linguists for centuries, computer scientists tend to use more and more linguistic – and especially semiotic – models to overcome the problem [2][3][4][5][6][7][8].

Multiagent systems and computational argumentation theory are fields of computer science where the disagreement over meaning issue is likely to occur. It has been shown that argumentation theory can allow different agents, which uses machine learning over different subsets of a training set, to match what they have learned using arguments [10]. Those arguments are features found in their environment or hypotheses about the rules of this environment. Initial implementations of the theory are confronted to the grounding problem when it comes to merging the results of a classification achieved by two agents on different subsets. Even in the case of supervised learning, there is a huge probability that the two agents use different generalizations. Those differences are due to the random factor when it comes to considering generalization as search in a graph space [9]. But the problem that we address is more general and can be applied to multiagent systems when the learning processes of the agents are different as well. This is because argumentation theory takes place *after* the learning, as an independent module [10].

Our long term research goal is to investigate how the solutions from semiotics and computational semiotics can meet the specific needs of argumentation theory. The answers provided will be used to build a communication protocol for multiagent systems. This protocol, based on argumentation theory, will use a semiotic representation of symbols [11]. This protocol should allow agents with different segregates of their environment to communicate – communication provided by the share of a same final set of symbols. In the section II, we justify our choice to use computational semiotics in order to reach our goal. We will then present a simple communication protocol in section III. This protocol is evaluated on three basic cases in section IV.

II. Motivation

A. Considerations about semiotics

The historical notion of symbol in computer science is clearly defined and has little to do with the notion of meaning [1]. Moreover, with this implementation of symbols intelligent systems are vulnerable to the symbol grounding issue mentioned in section I. This issue is easily overcome by human speakers [12][13]. They use their natural languages along with the ability to align those languages with each other. This observation on natural language speakers motivates us to transpose the human mechanisms of communication to the field of multiagent systems. By adopting a model of concepts which rely on a semiotic paradigm we are developing a communication protocol able to solve semantic alignment issues. This protocol does so by turning the semantic alignment problem into an agreement over the meaning problem, meaning that is not necessarily present in artificial intelligence, but that we can implement since the grounding problem is now considered as solved [1]. We need to introduce meaning in our protocol since it takes a central role in the linguistic description of human communication with natural languages [14][15]. A formal definition of the meaning is presented in the section III. For the moment we will simplify this definition by saying that meaning is what a listener considers that a symbol is standing for. Meaning is not anymore self-contained into the symbol, but emerges from the relation between the symbol and the speakers. Therefore, symbols' meaning is context-dependent – or at least social context dependent.

Our approach to model contextual meaning is based on the notion of *contrast set* developed in ethnographic studies of how people actually give meaning to words [16]. A contrast set is a collection of segregates, and a *segregate* is a "terminologically distinguished array of objects". For instance, a buyer can enter an eatery and ask "What kind of sandwiches ya got besides hamburgers and hot dogs?", to which the seller responds "How about a ham 'n cheese sandwich?". Here the collection of words describing the different kinds of products one can eat are the contrasts set: hamburgers, hot dog, ham 'n cheese sandwich, etc. However, the way one person segregates and the word or sign used to refer to them is contextually determined, which can lead to misunderstandings whose resolution requires agents' adaptation of their respective intended meanings. An example of misunderstanding (from [16]) is that of our client complaining by uttering this sentence: "Hey, that's no hamburger; that's a cheeseburger". The origin of the misunderstanding is that the client is considering *hamburger* and *cheeseburger* as two different elements in the contrast set he is using to conceptualize the eating options, while the seller considers that cheeseburger is the *subset by default* of the meal category hamburger.

The "meaning is use" paradigm is inspired by Wittgenstein's *Philosophical Investigations*, and is the underpinning of ethnographic notions of contrast set, as well as what Wittgenstein labels at "language games". We will approach the large problem of contextual meaning in a limited way, in scenarios where agents can negotiate and agree on meaning by building new contrast sets in a new context.

As a running example of context-dependent meaning we will use the common sense domain of *Furniture Shopping*. Let us assume they have some default meaning of some concepts (often called ontologies in Artificial Intelligence), for instance about furniture. If we ask two agents before they interact if an *armchair* is a chair they would probably answer affirmatively. For our purposes, we can set that *armchair* is a sub-concept of the *chair* concept. Now, imagine the buyer enters the shop and tells the seller: "I want to buy one armchair and four chairs". If the seller understands the meaning intended by the buyer – buying four chairs without arms and one with – no misunderstanding arises, and they will keep talking about "chairs" and referring to particular objects in the shop that are *chairs* without any disagreement on any specific object. Nevertheless, they are not using "chair" to refer as the same concept as before: now it means in fact *chairs without arms*. This is because the buyer has created the contrast set $\{$*armchair*, *chair*$\}$, and by doing so he has implicitly decided to use the word "chair" with a new intended meaning. If the two agents consistently use the term "chair" to refer only to objects in the shop that are chairs and are not armchairs, we say they have achieved an agreement on meaning. This shift in the meaning of a term or word is so pervasive that we humans are hardly aware of it, but we would consider it wrong if the seller tried to sell three armchairs and two chairs without arms (which is consistent with the default meaning of "chair" and "armchair").

The issue we need to address now is how to represent concept meaning in a way that allows us to have a computational model in which the sort of change in meaning illustrated in the furniture example is achieved by creating contrast sets. As we said, the approach is semiotic: a *concept* is represented by a semiotic triangle $\langle S, I, E\rangle$ with three components: a sign S, a meaning – its *intensional definition*, and an object or referent – its *extensional definition* [11]. In this view, a sign like "chair" can have two different meanings in the *Furniture shopping* scenario by belonging to two different semantic triangles. What we called the default meaning is often found in dictionaries and ontologies. It specifies the typical or more frequent sense of a sign like "chair", and could be expressed in a semiotic triangle $\langle$"chair"$, I, E\rangle$ where I is the default meaning of *chairs* (including *armchairs* and other sorts of *chairs*), and E is the set of objects that can be referred to by that sign. However, after the buyer introduces the new contrast set $\{$*armchair*,*chair*$\}$, the meaning of the sign "chair" has to change in order to achieve successful communication. In the *Furniture Shopping* scenario, the agreed meaning of that sign can be expressed in a new semiotic triangle $\langle$"chair"$, I', E'\rangle$, where now the agreed meaning I' is that of *chairs without arms* (because when referring to those the agents would use the "armchair" sign). Moreover, the set of objects that may stand as referents of the sign has also changed, since E' is about objects that are chairs but not armchairs.

Specifically, our computational model will assume two agents with possibly different contrast sets, and we will assume that each term in a contrast set is a sign S_1 represented as a semiotic triangle, in which the segregate corresponds to extensional definition E_1 of that triangle.

B. Related work

The present work expands the agent-based "concept convergence" approach in [**?**], where argumentation and agreement were not contextual but focused on a single, isolated concept. Besides the argumentation theory and machine learning background developed in that approach, we are listing a few other approaches related to our work below.

1) Computational Semantics: Attempts of using semiotics in computer science have already been done and led to the field of computational semantics. The aim of computational semantics is quite different than ours. The main objective of computational semantic is the representation of knowledge by one entity [4][5], not the communication between different agents. However, the mathematical theory of objects could be a useful source of inspiration for our future work, if we need to enhance the complexity of our interpretation of the semiotic triangle. The detailed structure of objects in computational semantics could be substituted to our simple features and therefor allows the agents to have more control over the construction of their meanings.

Computational semiotics starts from a will to make artificial intelligence similar to human intelligence by using the same semiotic approach of concepts [7][8]. On the contrary, our aim is to solve an identified issue from argumentation theory in computer communication by getting inspiration from natural language semiotic. We can oppose the two goals by stating that computer semiotics aims to be an exhaustive learning method of concepts while our approach is more a lazy learning method. The two approaches are working on the same field and can be seen as complementary, but they start from two different limits of this field.

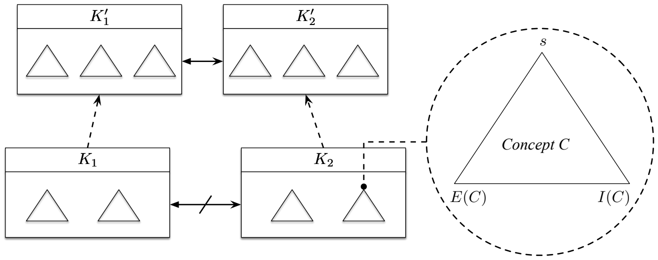

Fig. 1. A contrast set is a set of concepts that partition a domain of examples. To mutually adapt their meaning of concepts, each agent can create a new contrast set (in addition the old one) that allows it to reach an agreement over the meaning of a concept if that concept arose a disagreement in the old contrast set.

2) Ontology Matching: Ontology alignment has been studied on database schemas, XML schemas, taxonomies, formal languages, entity-relationship models, and dictionaries. Formally, while matching is the process of finding relationships or correspondences between entities of different ontologies, alignment is a set of correspondences between two (or more) ontologies (by analogy with molecular sequence alignment, according to [17]). Thus, alignment is the output of the matching process.

There are two families of approaches to ontology alignment: (1) syntactic approaches establish matchings among predicates, terms or other structural properties of a formalism, essentially focusing on a notion of similarity; (2) semantic approaches establish logical equivalence correspondences among ontology terms, essentially focusing on a notion of semantic equivalence. We considers here “semantic” in a logical paradigm. We propose a third approach, different of those two: a semiotic viewpoint that takes into account both the extensional and intensional definitions of a concept.

There is a second difference between ontology matching and an agent-based approach. For each agent, concept mapping is performed inside each individual agent, not by a third party comparing two (external) ontologies. Finally, we do not deal with an initial set of ontologies, as it is commonly done in ontology matching. We deal with agents that argue and agree over concepts based on how the concepts are used in a particular context.

III. A SEMIOTIC APPROACH TO MEANING ALIGMENT

Now that the issue of agreement over the meaning has been stated, we present our attempt to solve it. This attempt is motivated by our analysis of the state of the art presented in section II. The reader should keep in mind that the protocol is symmetric for both agent when he reads this section.

A. Concepts and Contrast Set

As we said, a concept C_i is understood as a semiotic triangle, i.e., as being composed of a sign s_i, an intensional definition $I(C_i)$ and an extensional definition $E(C_i)$. We use the notation $C_i = \langle s_i, I(C_i), E(C_i)\rangle$ to represent a concept C_i as shown in figure 1. $E(C_i)$ is understood as the set of objects or examples known by the agent as belonging to concept C_i. The intensional definition $I(C_i)$ contains a set $\alpha_1 \dots \alpha_m$ of generalizations such that $\forall e \in E(C_i), \exists \alpha \in I(C) : \alpha \sqsubseteq e$ and $\forall \alpha \in I(C), \exists e \in E(C_i) : \alpha \sqsubseteq e$. We will use the notation $I(C_i) \sqsubseteq e$ when there is a generalization in the intensional definition that subsumes the example e, and to say that an example belongs to the concept which sign is s_i we will also use the notation $s_i \sqsubseteq e$.

We introduced the notion of contrast set as a collection of concepts that induces a partition on a domain. We will now define a contrast set in which concepts are represented by the semiotic triangle. A contrast set K consists of a collection concepts $K = (C_1, \dots C_n)$, with a collection of signs $s(K) = (s_1, \dots, s_m)$, a collection of intensional definitions $I(K) = (I(C_i), \dots, I(C_m))$, and set of examples $E(K) = E(C_1) \cup \dots \cup E(C_m)$ belonging to those concepts. Since a contrast set determines a partition of the elements in $E(K)$, now each intensional definition $I(C_i) \in I(K)$ has to comply with the following property: $\forall e \in E(K) \setminus E(C_i) : I(C_i) \not\sqsubseteq e$ —i.e., the generalizations in $I(C_i)$ should not subsume an example also subsumed by a generalization from an intensional definition of another concept in the contrast set.

To explore the notion of agreement over concept meaning, we use a scenario with two agents that have individual (and possibly different) contrast sets over the same domain. The agents jointly observe new elements in this domain, and categorize elements in one of the concepts of their individual contrast sets. Disagreement arises when agents try to talk about a concept and the *signs* of the concept in which the example is categorized are different. Upon disagreement, the agents engage in an exchange of arguments to adapt their individual contrast sets to one another until the disagreement is solved. This is an iterative process in which both agents build two new contrast sets that are closer: $K_1 \xrightarrow{adapt} K_1' \rightleftarrows K_2' \xleftarrow{adapt} K_2$ (see Fig. 1 and Fig. 2).

B. Communication Protocol

We assume that our agents already share the language $\mathcal{L}$ and are able to exchange information through a communication protocol that we will specify, using messages with five different kinds of communicative acts:

- *Assert*(s, e): this message affirms that the sender considers e to be an example belonging to the same concept as the sign s ($s \sqsubseteq e$).
- *Accept*(s, e): this message confirms that the sender agrees on fact that e is in the concept whose sign is s ($s \sqsubseteq e$).
- *Refuse*(s, e): this message informs that the sender disagrees on fact that e is in the concept whose sign is s ($s \not\sqsubseteq e$).
- *Ask*(s, e): this message asks to the other agent which generalization, in the concept whose sign is s, subsumes e.
- *Answer*(s, β, e): sends the generalization β, in the concept whose sign is s, that subsumes e ($\beta \sqsubseteq e$).

The basic elements of our communication protocol have been presented. We will now present how an agent reacts to those communicative acts. The relation between the elements of language and the language user – in our case the agent – is central in Peircean Semiotics [11].

In order to explore the possibilities offered by the use of contrast sets, we implemented three different scenarios that we present in the section IV. We will now present the protocol used by two agents to build and organize their contrast sets in order to achieve a mutual agreement. A sketch of the protocol is as follows:

1) The two agents are waiting for the presentation of an example.
2) An example e_x is presented to the agents. Each agent A_i categorizes the example with a sign s_i and sends *Assert*(s_i, e_x) to the other agent.
3) Each agent verifies whether the received sign s_j is in its contrast set. If the sign is unknown the agent goes to step 5, otherwise they go to step 4.
4) Each agent verifies whether $s_j \sqsubseteq e_x$ in its individual contrast set:
   - if it is true, the agent does not disagree with the meaning of s_j $(s_i = s_j)$ and sends the message *Accept*$(s_{j,} e_x)$. Then the agent goes to step 1;
   - if it is not true, the two agents are going to create new concepts in order to reach an agreement over the contextual meaning of the sign used. They go to step 5.
5) If there is at least one agent which does not have one of the exchanged signs in its contrast sets, e.g., s_i, then this agent sends a message *Ask*(s_i, e_x) to the other agent.
6) The other agent sends back an *Answer*(s_i, α_m, e_x) where α_m is the generalization that subsumes e_x $(\alpha_m \sqsubseteq e)$.
7) The agent that did not know s_i uses the received generalization α_m to check if the unknown sign is a synonym or an hyponym of a known concept $I(C_j)$ (more details about the process of 7 are given at the end of the scenario):
   - in case of synonymy, the agent creates a new concept for the sign s_i;
   - in case of hyponymy, the agent creates two new concepts; one for s_i (the unknown sign) and one other for s_j (the sign this agent sent when categorizing e_x);
   - otherwise, the agent sends a *Refuse*(s_i, e_x).
8) The agent incorporates those concepts:
   - if no new contrast set has been created for the current social context yet, the agent creates one and puts the new concept(s) into it, while including as well the rest of the concepts from the first contrast set that are left unchanged by any changes;
   - if a contrast set has already been created, C_i is removed from it and the new concept(s) is/are added.
9) The agent returns to 3 (although this time it will not disagree at 4).

We will now provide more details about the internal process occurring during the step 7. When an agent receives the message *Answer*(s_i, α_m, e_x), its interpretation process starts by starts by creating the set of examples $E(C_i^*)$ that contains all the examples e_i^* from the extensional definition $E(K_l)$ of its current contrast set K_l which verify the property $\alpha_m \sqsubseteq e_i^*$. Then, $\forall C_j \in K_l$.

In the case of a synonym, the new concept created is: $C_a = \langle s_i, I(C_j), E(C_j) \cup e_x \rangle$. Since there is synonymy, there is no need to change the extensional and intensional definitions from the old concept C_j, except from adding the new example e_x to the extensional definition of course.

However, in the case of an hyponymy, the agent creates a second new extensional definition $E(C_j^{**}) = E(C_j) - E(C_i^*)$. At this point, the agent has created two new concepts, namely $C_a = \langle s_j, I(C_a), E(C_j^{**}) \rangle$ and $C_b = \langle s_i, I(C_b) = \alpha_m, E(C_i^*) \cup e_x \rangle$. After these two concepts are incorporated to the current contrast set it still need to generate the intensional definition of the C_a concept. Now, in order to create this intensional definition $I(C_a)$, we use the ABUI argumentation-based inductive learning method [**?**] that takes set of examples $E(K_n)$ an positive examples, and the rest of examples in the contrast set as negative examples.

- if $\exists C_j : E(C_j) = E(C_i^*)$, then the agent recognizes s_i as a synonym;
- if $\exists C_j : E(C_j) \subset E(C_i^*)$, the agent recognizes s_i as an hyponym (Fig. 2).

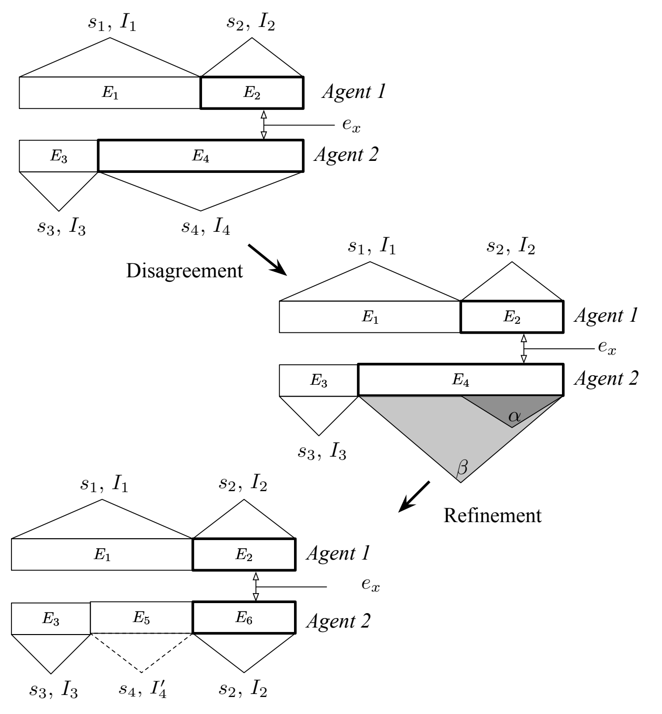

Fig. 2. When a new concept is needed, the agent asks the relevant generalization from the other agent's intentional definition of that concept; this generalization (e.g., α) leads to a split of the previous extensional definition E_4 in two: E_5 (examples subsumed by α) and E_4' (the rest).

IV. Scenarios

We present three scenarios to test the agents' ability to reach an agreement over some concept meaning. The examples

TABLE I. INITIAL CONTRAST SETS OF AGENTS A_1 AND A_2 IN CONCEPT HYPONYMY

C. set	Agent 1 contrast set K_1	
Concept	C_1	C_2
Sign	*stool*	*chair*
$I(C)$	α_1: *no.arm, no.back*	α_2: *with.back*

C. set	Agent 2 contrast set K_2	
Concept	C_3	C_4
Sign	*chair*	*armchair*
$I(C)$	β_1: *no.arm*	β_2: *with.arm, with.back*

TABLE II. FINAL CONTRAST SETS OF A_1 AND A_2 IN CONCEPT HYPONYMY

C. set	Agent 1 contrast set K_1'		
Concept	C_1	C_5	C_6
Sign	*stool*	*chair*	*armchair*
$I(C)$	α_1: *no.arm, no.back*	α_3: *no.arm, with.back*	β_2: *with.arm, with.back*

C. set	Agent 2 contrast set K_2'		
Concept	C_7	C_8	C_4
Sign	*stool*	*chair*	*armchair*
$I(C)$	α_1: *no.arm, no.back*	β_4: *no.arm, with.back*	β_2: *with.arm, with.back*

used in those scenarios are referring to the common sense domain of *seats*. Specifically, examples of seats are divided in three categories: Chairs, Armchairs and Stools, where we are using their typical definitions: (1) a stool has no back and no arms, (2) a chair has a back but no arms, and (3) an armchair has a back and arms. In every scenario, agents A_1 and A_2 start with their own individual contrast set. Given a new example e_x, they will try to solve their disagreement when it occurs.

A. Concept Hyponymy

In the first scenario the two agents have initially their contrast sets K_1 and K_2, shown in Table I. The first example e_1 presented to both agents is an armchair. The two agents first send *Assert*(*chair*, e_1) in the case of A_1 and *Assert*(*armchair*, e_1) in the case of A_2. They both discover a disagreement. The sign *armchair* $\notin S(K_1)$ so A_1 sends *Ask*(*armchair* , e_1) to A_2. A_2 responds with *Answer*(*armchair* β_2) since $\beta_2 \sqsubset e_1$. Then, A_1 creates the subset $E(C_6) = \{e_i \in E(C_2) | \beta_2 \sqsubseteq e_i\}$. Since $E(C_6) \neq E(C_2)$, A_1 creates the subset $E(C_5) = E(C_2) - E(C_6)$. No contrast set has been created yet, so A_1 creates a new contrast set K_1'. The new concepts $C_5 = \langle \textit{chair}, I(C_5), E(C_5)\rangle$ and $C_6 = \langle \textit{armchair}, \beta_2, E(C_6) \cup e_1\rangle$ are added to K_1' with the concept C_1. Then, A_1 performs an induction on the new $E(K_1')$ for the sign *chair* that results to the generalization α_3 which is added to $I(C_5)$. Finally, A_1 sends an *Accept*(*armchair*).

The second example presented, e_2, is a chair. The agents send to each other *Assert*(*chair*, e_2). Since both agents agree on the use of *chair* as a sign to describe a chair, they just send to each other two messages *Accept*(*chair*, e_2).

The last example e_3 presented to the agents is a stool. As with e_1, they both notice the disagreement over *stool* and *chair*, but this time *stool* $\notin K_2$. It leads A_2 to send *Ask*(*stool*, e_3) to A_1. Since $\alpha_1 \sqsubset e_3$, A_1 sends back *Answer*(*stool*, α_1). Now it is A_2's turn to create a subset $E(C_7) = \{e_i \in E(C_3) | \alpha_1 \sqsubseteq e_i\}$, and since $E(C_7) \neq E(C_3)$, A_2 creates the subset $E(C_8) = E(C_3) - E(C_7)$. No new contrast set has been created by A_2, so it creates K_2'. The new concepts $C_7 = \langle \textit{stool}, \alpha_1, E(C_7) \cup e_3\rangle$ and $C_8 = \langle \textit{chair}, I(C_8), E(C_8)\rangle$ are added to K_2' with the concept C_4. A_2 performs an induction on the new $E(K_2')$ for the sign *chair*. The resulting generalization β_4 is added to $I(C_8)$. Finally, A_2 sends an $Accept$(*stool*, e_3) to A_1. The contrast sets are now as shown in Table II. We can now see that K_1' and K_2' are mutually intelligible.

B. Concept Synonymy

TABLE III. A_1 AND A_2 CONTRAST SETS IN CONCEPT SYNONYMY

C. set	Agent 1 contrast set K_1		
Concept	C_1	C_2	C_3
Sign	*stool*	*chair*	*armchair*
$I(C)$	α_1: *no.arm, no.back*	α_2: *no.arm, with.back*	α_3: *with.arm, with.back*

C. set	Agent 2 contrast set K_2		
Concept	C_4	C_5	C_6
Sign	*stool*	*chair*	*recliner*
$I(C)$	β_1: *no.arm, no.back*	β_2: *no.arm, with.back*	β_3: *with.arm, with.back*

In the second scenario the two agents have initially their contrast sets K_1 and K_2 as shown in Table III. The only example e_1 presented to both agents is an armchair. A_1 sends *Assert*(*armchair*, e_1) to A_2 and A_2 sends *Assert*(*recliner*, e_1) to A_1. It leads to a disagreement between A_1 and A_2. Since *armchair* $\notin K_2$ and *recliner* $\notin A_1$ neither, both agents can send an *Ask* message. Let's say that A_1 receives the *Ask* message first; A_1 will respond with *Answer*(*armchair*, α_3), where α_3 is the generalization that was used by A_1 to categorize e_1. A_2 creates a new set of examples $E(C_7) = \{e_i : e_i \in E(C_6) | \alpha_3 \sqsubseteq e_i\}$. No contrast set has been created before so A_2 creates K_2'. Since $E(C_7) = E(C_6)$, a new concept $C_7 = \langle \textit{armchair}, I(C_6), E(C_6) \cup e_1\rangle$ is created and added to K_2' along with C_4 and C_5. *Accept*(*armchair*,e_1) is sent to A_1 by A_2. We notice that if A_1 had been faster than A_2 to send its Ask message, it is *recliner* that would have been used to designate Armchair by both agents. This would not have hindered them from reaching mutual intelligibility.

C. Concept Teaching

This time K_1 is still as it was in Table III but K_2 is modified as shown in Table IV. The first example e_1 presented to the agents is a stool. A_1 sends *Assert*(*stool*, e_1) to A_2 and A_2 *Assert*(*seat*, e_1) to A_1. Since *seat* $\notin S(K_1)$ and *stool* $\notin S(K_2)$, each agent can send an *Ask* message to the other. Let's say that A_1 is the fastest to send its *Ask*(*seat* message, e_1), A_2 sends *Answer*(*seat*, e_1).$any \sqsubset \alpha_1, \alpha_2$ and α_3: A_1 sends *Refuse*(*seat,* e_1) to A_2. Now A_2 sends *Ask*(*stool*, e_1) to A_1. A_1 sends back *Answer*(*stool*, α_1). Now A_2 creates the subset $E(C_5) = \{e_i : e_i \in E(C_4) | \alpha_1 \sqsubseteq e_i\}$, and since $E(C_5) \neq E(C_4)$, A_2 also creates the subset $E(C_6) = E(C_4) \setminus E(C_5)$. No contrast set has been created yet so A_2 creates K_2'.The new concepts $C_5 = \langle \textit{stool} , \alpha_1, E(C_5) \cup e_1\rangle$ and $C_6 = \langle \textit{seat}, I(C_6), E(C_6)\rangle$ are added to K_2'. New generalizations for *seat*, β_3 = *with.back* and β_4 = *with.arm*, are learned by induction from $E(K_2')$ and added to $I(C_6)$.

The second example presented e_2 is a chair. A_1 sends *Assert*(*chair*, e_2) to A_2 and A_2 *Assert*(*seat*, e_2) to A_1. Again, *seat* $\notin S(K_1)$ and *chair* $\notin S(K_2')$ so each agent can send an *Ask* message to the other. Let's say that this time A_2 is faster and sends *Ask*(*chair*, e_2) to A_1. A_1 sends back *Answer*(*chair*, α_2) since $\alpha_2 \sqsubset e_2$. A_2 creates a new set $E(C_7) = \{\forall e_i :$

TABLE IV. A_2'S CONTRAST SETS BEFORE CONCEPT TEACHING

C. set	**Agent 2 contrast set** K_2
Concept	C_4
Sign	*seat*
$I(C)$	β_1 : *any*

TABLE V. A_1 AND A_2 CONTRAST SETS AFTER THE LAST CONCEPT LEARNING IN CONCEPT TEACHING

C. set	**Agent 2 contrast set** K'_2		
Concept	C_5	C_7	C_9
Sign	*stool*	*chair*	*armchair*
$I(C)$	β_2: *no.arm, no.back*	α_2: *no.arm, with.back*	β_6: *with.arm, with.back*

$e_i \in E(C_6)|\alpha_2 \sqsubseteq e_i\}$, and since $E(C_7) \neq E(C_6)$, A_2 also creates the subset $E(C_8) = E(C_6) \setminus E(C_7)$. The new concepts $C_7 = \langle$*chair* $, \alpha_2, E(C_7) \cup e_2\rangle$ and $C_8 = \langle$*seat* $, I(C_8), E(C_8)\rangle$ are added to K'_2. The concept C_6 is removed from K'_2. A generalization for *seat*, β_6, is learned by induction from $E(K'_2)$ and put into $I(C_8)$. A_2 sends *Accept*(*chair*, e_2) to A_1.

The last example presented e_3 is an armchair. A_1 sends *Asset*(*armchair*, e_3) to A_2. Meanwhile A_2 sends *Assert*(*seat*, $_3$) to A_1. Since $seat \notin S(K_1)$ and $armchair \notin S(K'_2)$, both agents can send an *Ask* and let's say that this time again A_2 is quicker and sends *Ask*(*armchair*, e_3) to A_2. Since $\alpha_3 \sqsubset e_3$, A_1 sends back *Answer*(*armchair*, α_3). After receiving the answer, A_2 creates a new set of examples $E(C_9) = \{\forall e_i : e_i \in E(C_8)|\alpha_3 \sqsubseteq e_i\}$. Since $E(C_9) = E(C_8)$, a new concept $C_9 = \langle$*armchair*, $I(C_8), E(C_8) \cup e_3\rangle$ is created and added to K'_2 while C_8 is removed. An *Accept*(*armchair*,e_3) is sent to A_1. Table V shows the new K'_2.

V. DISCUSSION

The three scenarios discussed in IV show that agreement over meaning can solve semantic alignment issues in a simple way. Two agent do not have to match all their representations – whatever they are, here concepts – in order to communicate. Since our agents do not try to be exhaustive, but rather pragmatic in their alignment, we can qualify our approach as a lazy method. However, with this lazy method, it is possible to reach an agreement on the use of a set of symbols. Those symbols, therefore, are designed and implemented in a semiotic conception. This conception allows them to be flexible enough to achieve an online refinement of the meaning of the signs they use.

An important point of this protocol design is its cooperation-oriented mechanism. Argumentation theory can be seen as a competitive theory, with each agent trying to attack other agents' arguments and defend its own. However, in our protocol the agents try to find how to adapt their symbols' representations in order to agree with their interlocutor. This becomes possible since agents are not in a Realist paradigm where they have to defend their beliefs that are considered as truths. They are the only one in charge of what is true and what is not. Therefore, they have a pragmatic approach and their goal is to be able to communicate with each other in a minimal number of steps.

It is possible to argue about the usefulness of extensional definitions in our model. In fact, the model could be adapted to work only with the signs and the intensional definitions. The agent's test to determine if a received intensional definition is a hyponym or a synonym of one of its own sign can be done directly on the intensional definition. According to this observation, extensional definitions only bring more complexity in the model by eventually necessitating to be split. In fact, we plan to give the extensional definition a central role in the future iterations of our model. This idea is explained in section Future Work.

Another arguable aspect in the conception of our model is the presence of semantic content in message types (*Assert, Ask, Answer, Accept* and *Refuse*). The presence of performative information in the structure of a message is common in multiagent systems protocols, for example in FIPA [18]. But in our case, we want a fully context dependent meaning. Since performative information can be considered as the structure of our protocol, it is hard to imagine how to get rid of it. However, recent studies on multiagent systems composed of agents able to communicate without sharing the same protocol could lead to a solution for this problem.

Another objection along similar lines is that our system, our system uses a set of already-grounded symbols. The features of our examples are interpreted in the same way by both agents – they are sharing the language $\mathcal{L}$ – and it means that for the moment, we are basically moving the symbol grounding issue to another level of detail in the meaning, but we are not addressing the question of the meaning *itself*. However, we can consider this as a recursive problem. Therefore, it becomes a graph theory problem where all the level of details can be discussed between the two agents. This solution is detailed in the section V.

In the scenarios, we have not tested the performances of our model against other computational solutions presented in the state of the art. We believe that it would make little sense at this early stage of our research. A first point is that we have no ontologies in our model. If we wanted to test it against ontology matching algorithms, we would have to provide them an ontology of the problem or to make the agents able to learn those ontologies. It means that we should either provide information to those algorithms, information that our model does not need, or add a learning phase and study its impact. However, we stated that a key point of our model is its learning-independent feature. To compare it to other existing approaches to the problem, there is an issue with complexity.. It is complicated to propose a valid test to evaluate our model on such simple scenarios. New iterations of the protocol should introduce more complexity and allow us to actually compare the performances of our model to other state-of-the-art models.

We close this discussion with a review of our model's limitations – as announced in section III. Our model can only handle cases where two concepts in conflict have a hyponymy or hyperonymy relation. In such cases, the attitude to adopt is obvious: the agent confronted to a hyponym of one of its concepts has to learn this hyponym. Doing nothing would not solve the disagreement and making the agent learning only hyperonyms would lead to build more general meaning of concepts at each disagreement, which would lead to a loss of information over time. But when a concept's intensional definition from an agent A_1 subsumes a set of examples belonging to two different extensional definitions of another agent A_2, the situation is reciprocal. In this case, A_2 would also have one of its intensional definitions subsuming a set of examples belonging to two different extensional definitions of A_1: which agent should change its intensional definition? We call this the "overlapping definitions" problem. In section Future Work, we propose our ideas to calculate the cost of each possible change in this situation, and to take an optimal decision.

Future Work

The extensional definition is a central element of our future work. We saw that one of our model's limitation is the need of some grounded symbols. Those symbols have a shared meaning for the two agents. The agents can disagree on the meaning of *chair* or *armchair* but to overcome this disagreement they need: (1) signs to describe the elements of the lexical domain – the features – as *with.arm*, *with.back* or *without.back* and (2) those signs – whichever they are – have to be numerous enough to create a partition of this lexical domain. The second condition is not too difficult to fulfill in the simple scenarios that we are investigating. The first point is a more problematic issue. There is no particular reason why the sign *with.arm* should be different from the sign *arm*. Each sign should be described by a combination of feature terms and each feature term should be considered as a sign. This leads us to a recursive approach where agents can have several levels of contrast sets and are able to argue about the meaning of all the signs of their language $\mathcal{L}$. We do not want to address the specific issue of the learning of a set of new signs, we rather investigate how the systemic relations of all the elements from the already existing $\mathcal{L}$ can dynamically change in order to reach an agreement. We plan to consider the examples of our extensional definition as possible intensional definitions for more specific concepts to overcome this issue. In our scenario, the extensional definition of *seat* would be the intensional definition of *chair*, *armchair* and *stool*. However, we are still in the particular case. The relations between the concepts are just hyponymy and hyperonymy. With overlapping definitions the process would be more difficult to design since the relations between concepts would not draw an oriented graph anymore.

The main issue with our model in its current stage of development is the overlapping definitions problem. It is necessary to evaluate the cost of a conceptual shift to be able to decide which agent should change its concept. Our first thought was to take the number of examples from the extensional definition subsumed by the new intensional definition as a cost function. However, since we are trying to set the elements of our extensional definition as intensional definitions of sub-concepts, this cost function would be susceptible to change over the time. It is a problem since an optimal decision at t might become the less optimal solution at $t+1$ for the same situation. This would mean that we have to re-evaluate all the concepts of a contrast set every time than an intensional definition is modified – independently of its position in the global hierarchy of contrast sets. We have to evaluate if this solution would be better or worse than an arbitrary decision. We consider here the best and worst in terms of number of changes operated on a long scale of time. The naive answer would be that this solution should lead at least to the same results as an arbitrary decision since it is optimal on a one-trial scale. But a large number of examples can also be an index of a too general concept that will have to be changed anyway on a long run. In this case, it could be more useful to see the number of examples as a gain function rather than a cost function.

The next step of our work is to introduce more complexity in our lexical domains. We started with only three words and a dozen features – with only three of them being relevant to determine a sign's use in the most complicated case. In order to compare our model with other state-of-the-art models, we would like to handle more complex examples. In order to do so we plan to take examples from the biological domain of marine sponges. The associated data set has been already used in argumentation theory studies [19]. Since its taxonomy is still debated by researchers in phylogeny, we can assume that the complexity of the task is high enough to offer a good material for tests.

For the moment we limited our scenarios to two-agent systems. We are not interested in testing our model with more than two agents at this point of our research. This change would introduce specific issues by complicating drastically the rules of speak priority. We want to face this issue only after having already built a fully functional model working with two agents. The case of diachronic evolution of used signs in agent networks has been already studied [20], but not with a similar protocol as ours. However, the results of this study on the dynamic of words' use provides a solid ground to have a hint of what we should expect from this scenario.

We have fulfilled the objectives that we fixed to ourselves for this first investigation. Moreover, we have now serious trails on how to continue and complete them.

Acknowledgements: This paper has been partially supported by projects ESSENCE: Evolution of Shared Semantics in Computational Environments (ITN 607062) and NASAID (CSIC Intramural 201550E022)

References

[1] L. Steels, "The symbol grounding problem has been solved. so what's next," in *Symbols and Embodiment: Debates on Meaning and Cognition*. Oxford University Press, 2008, pp. 223–244.

[2] J. F. Sowa, "Ontology, metadata, and semiotics," in *Conceptual structures: Logical, linguistic, and computational issues*. Springer, 2000, pp. 55–81.

[3] S. Regoczei and G. Hirst, "The meaning triangle as a tool for the acquisition of abstract, conceptual knowledge," *International journal of man-machine studies*, vol. 33, no. 5, pp. 505–520, 1990.

[4] R. Gudwin and F. Gomide, "An approach to computational semiotics," *NIST SPECIAL PUBLICATION SP*, pp. 467–470, 1997.

[5] B. B. Rieger, "Computational semiotics and fuzzy linguistics," in *Proc. of the 1997 International Conference on Intelligent Systems and Semiotics*. Citeseer, 1997.

[6] D.-H. Gerd, "Humanlike computational learning theory : A computational semiotics perspective," in *AFRICON, 2009*. IEEE, 2009, pp. 1–6.

[7] J. A. Guerrero, A. S. Gomes, and R. R. Gudwin, "A computational tool to model intelligent systems," *Anais do 4o SBAI-Simpósio Brasileiro de Automação Inteligente*, pp. 8–10, 1999.

[8] A. M. Meystel, "Intelligent systems: A semiotic perspective," in *Intelligent Control, 1996., Proceedings of the 1996 IEEE International Symposium on*. IEEE, 1996, pp. 61–67.

[9] T. M. Mitchell, "Generalization as search," *Artificial intelligence*, vol. 18, no. 2, pp. 203–226, 1982.

[10] S. Ontañón and E. Plaza, "Similarity measures over refinement graphs," *Machine Learning*, vol. 87, no. 1, pp. 57–92, Apr. 2012.

[11] J. Queiroz and F. Merrell, "On peirce's pragmatic notion of semiosis : A contribution for the design of meaning machines," *Minds and Machines*, vol. 19, no. 1, pp. 129–143, 2009.

[12] B. Galantucci, "An experimental study of the emergence of human communication systems," *Cognitive science*, vol. 29, no. 5, pp. 737–767, 2005.

[13] M. J. Pickering and S. Garrod, "Toward a mechanistic psychology of dialogue," *Behavioral and brain sciences*, vol. 27, no. 02, pp. 169–190, 2004.

[14] A. Cangelosi and S. Harnad, "The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories," *Evolution of communication*, vol. 4, no. 1, pp. 117–142, 2001.

[15] H. H. Clark and S. E. Brennan, "Grounding in communication," in *Perspectives on Socially Shared Cognition*. American Psychological Association, 1991, pp. 13–1991.

[16] C. O. Frake, "The ethnographic study of cognitive systems," in *Language, Culture, and Society*, B. G. Blount, Ed. Waveland Press, 1969, pp. 125–136.

[17] J. Euzenat and P. Shvaiko, *Ontology matching*. Springer-Verlag, Heidelberg, 2007.

[18] FIPA, "FIPA ACL message structure specification," FIPA Agent Communication Language Specifications, 2002. [Online]. Available: http://www.fipa.org/specs/fipa00061

[19] S. Ontañón and E. Plaza, "Concept convergence in empirical domains," 2010, pp. 281 – 295.

[20] G. Schaden, "Markedness, frequency and lexical change in unstable environments," in *Formal and Experimental Pragmatics*, M. F. J. Degen and N. Goodman, Eds., Germany, 2014, pp. 43–50. [Online]. Available: http://hal.univ-lille3.fr/hal-01073128