Robust Semantic Interpretation and Dialog Management in the Context of a CALL Application

Johan Michel
Avenue Laënnec
72085 LE MANS CEDEX 9
johan.michel@univ-lemans.fr

Jérôme Lehuen
Avenue Laënnec
72085 LE MANS CEDEX 9
jerome.lehuen@univ-lemans.fr

Abstract

This paper describes a work about dialog managing in the context of a Computer Assisted Language Learning (CALL) research. In this paper, we choose to focus on the dialog management which is modeled in terms of tasks and methods. The following sections describe the semantic analyzer, the dialog model, and a few results.

1 Introduction

We will present a double model (semantic interpretation and dialogue management) for human-computer dialog in the context of a computer-assisted language learning (CALL) system. In this environment, the learner is implicated in an interaction with a virtual partner around a task (a recipe) to perform in a virtual micro-world (a virtual kitchen) [Lehuen 00], [Michel & Lehuen 02]. Related works exist, we can mention [Hamburger 94]. The following figure shows the user interface of our system:

2 The analyzer

In this specific context, the system will have to deal with incomplete or ungrammatical utterances. So our constraints are the following ones:

- Modularity: it takes place in an existing CALL architecture. More particularly, it is in connection with a virtual environment implemented as a micro-world;
- Robustness: it has to deal with odd utterances. Moreover, even if the learner uses correct words which are not in the lexicon, the interaction has to be continued;
- Non-determinism: it has to be able to produce partial or multiple interpretations for one utterance in order to carry on the interaction. The context of the interaction has to complete or select them.

We start from existing robust methods like "skimming parsing" [Dejong 82] and "chart parsing" [Winograd 83]. But the semantic analyzer we implemented is able to generate lexical hypothesis when unknown words impede its process [Michel & Lehuen 04]. Then, these hypotheses are used to engage a dialogic recovery strategy using the words recognized by the analyzer.

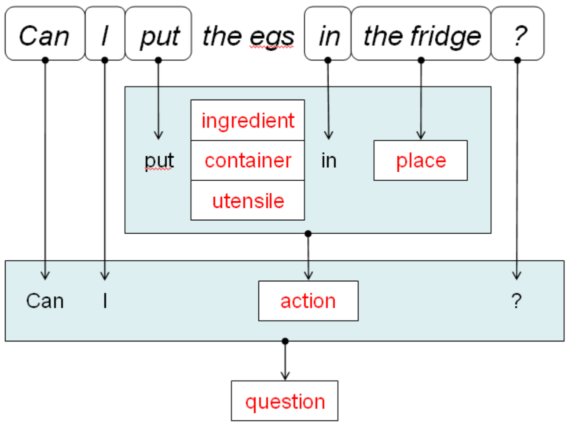


Fig. 1: Example of a syntax-driven hypothetico-deductive analysis

The figure 1 shows how the sentence *"Can I put the egs in the fridge?"* is analyzed as a question on the basis of one (triple) hypothesis about the unknown (and wrong) segment *"the egs"*. In this syntax-driven hypothetico-deductive analysis, *"the egs"* can be an ingredient, a container or a utensil. The analysis has three different steps: the lexical cover checking, the syntactic cover checking and the syntactic recovery. The lexical cover checking verifies if all the words in the utterance belong to the lexicon. The syntactic cover checking verifies if a syntactic pattern can be applied to the utterance. The syntactic recovery reorganizes the utterance to find a syntactic pattern. This recovery succeeds if after having reorganized the sentence, a syntactic pattern is found.

<table>
<tr><td rowspan="4">Good lexical cover</td><td>good syntactic cover</td><td colspan="2">« put the eggs in the fridge » (1)</td></tr>
<tr><td rowspan="3">bad syntactic cover</td><td>Syntactic recovery succeeds</td><td>« put in the fridge the eggs » (2)</td></tr>
<tr><td rowspan="2">Syntactic recovery does not succeed</td><td>« the fridge » (3)</td></tr>
<tr><td>« open » (3)</td></tr>
<tr><td rowspan="5">Bad lexical cover</td><td rowspan="2">good syntactic Cover</td><td>With hypothesis</td><td>« foo put the eggs in the fridge » (4)</td></tr>
<tr><td>Without Hypothesis</td><td>« put the foo in the fridge » (5)</td></tr>
<tr><td rowspan="3">bad syntactic cover</td><td>Syntactic recovery Succeeds</td><td>« put in the frige foo the eggs » (6)</td></tr>
<tr><td rowspan="2">Syntactic recovery does not succeed</td><td>« foo the eggs foo » (7)</td></tr>
<tr><td>« foo » (8)</td></tr>
</table>

Table 1. Examples of cases the analyser must handle

Different interaction strategies can be chosen given the analysis results. The first step is checking if the analysis's results correspond to the applicative (state of the applicative task) and interactive context, in this case, the analysis is validated. The second step is creating the partner's reaction from the applicative and interactive context.

3 The dialog model

The second part of our work focus on the knowledge of the dialog and the domain levels which are both modeled as tasks and methods. This approach is coming from research on generic mechanisms for problem-solving. It enables to rationalize the behavior of the system and provides a framework to design an abstract, implementation-independent description of problem-solving process (fig. 2). In our case, the domain level is only a pretext to engage dialog situations and to make rise linguistic problems. So, the dialog level is weakly connected with the domain level: the repair strategies are more about language and less about the task going on.

Figure 3 contains some tasks and methods to engage a dialog. You can see four tasks, one decomposition method, one iteration method and two execution methods.

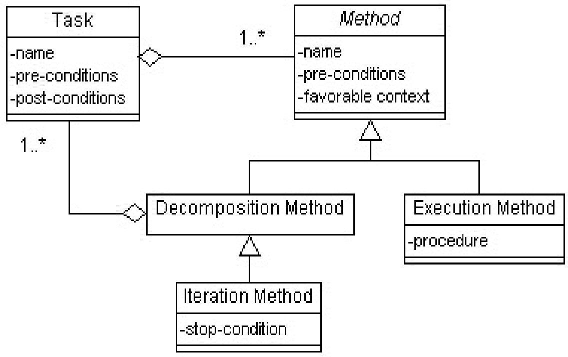


Fig.2: Task-Method framework as an UML class-diagram

.

```
(task (name     T-structure-dialog)
      (methods  M-structure-dialog))

(method-decomp (name   M-structure-dialog)
               (tasks  T-open-dialog T-dialog T-close-dialog))
```

```
(task (name     T-open-dialog)
      (methods  M-open-dialog))

(method-exec (name      M-open-dialog)
             (function  F-open-dialog))

(deffunction MAIN::F-open-dialog ()
   (assert (to-write "Hello, can you explain to me how to make a
                     chocolate cake?"))
```

```
(task (name     T-dialog)
      (methods  M-dialog))

(method-iter (name       M-dialog)
             (stop-cond  "(stop dialog)")
             (tasks      T-listen T-respond))
```

```
(task (name     T-listen)
      (methods  M-listen))

(method-exec (name      M-listen)
             (function  F-listen))

(deffunction MAIN::F-listen ()
   (assert (waiting-learner-utterance))
   (focus IHM))
```

Fig. 3: Some tasks and methods to engage a dialog

According to [Trichet & Tchnounikine 99], a Task-Method model can be decomposed into four layers, each layer manipulating data from the lower one. The highest layer, named control, deals with the way the modelized task is performed. It can be a simple algorithm, a classical expert system, or even a higher-level Task-Method engine (fig. 4). In our case, the control of the dialog level is performed by a simple independent algorithm, whereas the control of the domain level is performed by specific methods of the dialog level.

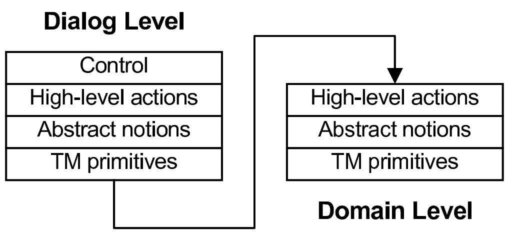


Fig.4: Task-Method model based on 4 layers

These models are fully implemented in Jess, a rule-based environment for the Java platforms. A first experiment has been conducted to test the system and its robustness. In the interaction logs, we have observed the different awaited dialog situations.

4 Examples of dialogs

In the logs, we found the different interaction cases we expected. For a given group, on forty-six exchanges, we counted thirty-two exchanges linked to the task without any problems (Figure 5), fourteen exchanges with linguistic problems (Figure 6), and one exchange with a problem about the task (Figure 7). The exchanges are translated from French.

```
Learner> Take the milk
Partner> I take the milk from the fridge.
Learner> Take the cream
Partner> I take the cream from the fridge.
```

Fig. 5: Exchanges without problems

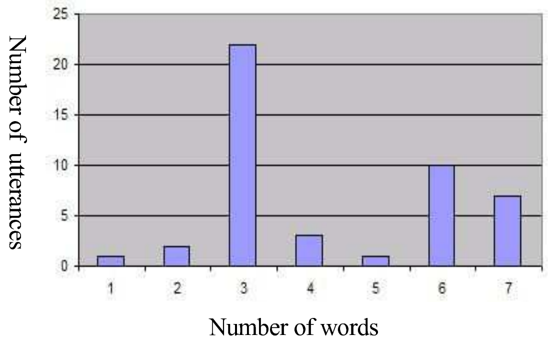
```
Learner> Open the cupboard at the top on
the righ
Partner> I do not understand when you write
« Open the cupboard at the top on the
righ». Which cupboard do you want to open?
Learner> Open the cupboard at the top on
the right
Partner> I open the cupboard.
```

Fig.6: Exchanges with linguistic problems

```
Learner> Take the milk
Partner> I understand, but it has already
been done
```

Fig.7: Exchange with a problem about the task

The learner's utterances were relatively short (from one to seven words) but this is mainly due to the nature of the task. The figure below illustrates this phenomenon for a given dyad:



References

[Dejong 82] G. Dejong 1982, *An overview of the FRUMP system*, In: "Strategies for natural language processing", Lehnert & Ringle, Erlbaum, Hillsdale, pp. 149-176.

[Delorme & Lehuen 03] F. Delorme and J. Lehuen 2003, *Dialog Planning and Domain Knowledge Modeled in terms of Tasks and Methods - A Flexible Framework for Dialog Managing*, In: ISMIS'03, Maebashi City (Japan), Springer-Verlag "Foundations of Intelligent Systems", Lecture Notes in Computer Science, pp. 689-693.

[Hamburger 94] H. Hamburger 1994, *Foreign language immersion: Science, practice, and a System*, In: JAIED Special Issue on Language Learning, vol. 5(4) pp. 429-453.

[Lehuen 00] J. Lehuen 2000, *A Dialogue-Based Architecture for Computer Aided Language Learning*, In: AAAI Fall Symposium on "Building Dialog Systems for Tutorial Applications", pp. 20-27.

[Lehuen & Luzzati 99] J. Lehuen and D. Luzzati 1999, *Acquisition coopérative d'une compétence langagière interprétative en dialogue homme-machine*, In: TALN'99, Cargèse (Corse), pp. 357-362.

[Michel & Lehuen 04] J. Michel and J. Lehuen 2004, *Un analyseur hypothético-déductif non déterministe pour l'apprentissage et la pratique d'une langue*, In: TAL & Apprentissage des Langues, Grenoble (France), pp. 13-22.

[Michel & Lehuen 02] J. Michel and J. Lehuen 2002, *Conception of a Language Learning Environment based on the Communicative and Actional Approaches*, In: ITS'2002, Biarritz (France), pp. 651-660.

[Trichet & Tchounikine 99] F. Trichet and P. Tchounikine 1999, *DSTM: a Framework to operationalise and Refine a Problem-Solving Method modeled in terms of Tasks and Methods*, In: International Journal of Expert Systems with Applications, vol. 16, pp.105-120.

[Winograd 83] T. Winograd 1983, *Language as a Cognitive Process*, Syntax, Addison-Wesley Publishing Compagny.

Extensions to Speaker/ Hearer Representation in DRT

Yafa Al-Raheb
University of East Anglia
y.al-raheb@uea.ac.uk

1 Introduction

Two DRT representations are introduced in Kamp et al. (2005). The first of which deals with presupposition and the second with propositional attitudes. However, neither representation deals with degrees of belief nor with speaker/hearer representation. The paper proposes a reconciliation and extension to these two DRT variants in order to represent degrees of belief and enhance the link between the linguistic content (utterance) and speaker/hearer representation, thus achieving an enriched account of presupposition.

2 Examples

The 'strength' of beliefs held by speakers differs from one situation to another, and depends on whether the speaker is introducing the topic of the dialogue. A weaker form of belief, called acceptance, is introduced to permit a form of differing degrees of beliefs. Acceptance represents the grey area where information is put on hold, not yet believed, but not rejected (Al-Raheb 2004). To explain what is meant by belief and acceptance, here is an example:

(1)

S1: I must buy Vincent's wife a birthday present.

H1:I didn't know Vincent was married.

S2: Yes, he is. His wife likes chocolate.

H2: She may also like flowers.

S3: I'll buy her chocolates.

The speaker, S, presents the presupposition (here it is new information to the hearer, H) that Vincent has a wife. Initiating the topic of presupposition allows H to attribute a stronger degree of belief to S about the presupposition. If we contrast example 1 with example 2 below, the stakes of the strength of beliefs would be much higher for H when he is required to perform an action than when simply going along with the dialogue.

(2)

S1: You should buy Vincent's wife a birthday present.

H1: I didn't know Vincent was married.

S2: Yes he is. His wife likes chocolate.

H2: She may also like flowers.

S3:But she prefers chocolate.

H3: I'll get her some chocolate.

In example 1, where H was not required to perform an action, it is safer for S to assume that H accepts the presupposition, as H is not committing to doing any task, than to assume the stronger case, i.e. H believes the presupposition. However, in example 2, where H agrees to buy Vincent's wife a present in H3, i.e. H commits to perform an action for Vincent's wife, S concludes that H believes the presupposition and adds this to S's representation of H's beliefs. These two examples show that certain pragmatic conditions can have a bearing on 'strength' of beliefs, which go beyond truth conditions. Our focus is on the compatible representation of 'strength' of be-

Proceedings of the 9th Workshop on the Semantics and Pragmatics of Dialogue, June 9-11, 2005, Nancy, France.

lief arising in presupposition.

3 Speaker/ Hearer Representation

This section explores the relationship between Kamp et al.'s two DRT variants, which can be linked and extended to provide appropriate representation for Speaker/ Hearer interaction. Kamp et al. (2005) discuss two separate variants of DRSs (Discourse Representation Structures) for beliefs and presupposition in DRT. The first only includes presuppositional and non-presuppositional aspects such as those shown in figure 1, to represent the linguistic content of an utterance without mention of cognitive states. Figure 1 shows Kamp's separation of presupposition and assertion for 'The rabbit is white' without dealing with beliefs. The first two nested DRSs show presuppositional information. The third nested DRS contains the non-presuppositional information, the assertion.

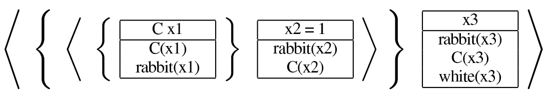


Figure 1: Linguistic Content DRS

The second DRS variant deals with beliefs, desires and intentions but not the presuppositional and non-presuppositional content. Figure 2 shows Kamp et al.'s (2005) formulation of someone who is shown a box full of stamps and told he can keep one. That person sees part of a stamp that he deems valuable, so he forms the belief, BEL, that there is a valuable stamp in the box, has the desire, DES, to possess the stamp, and the intention, INT, to pick the stamp from the box to fulfill his desire to own the stamp, 2d1840GB.

However Kamp et al. (2005) do not establish the link between these two DRT variants to explain the connection between speaker generation, speaker's utterance, and hearer recognition. To establish this link, a new

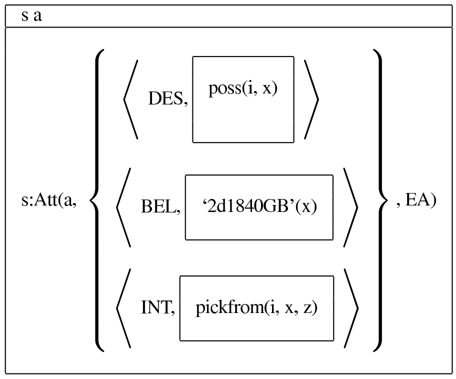


Figure 2: Propositional Attitudes DRS

DRS representing both the linguistic content and cognitive states of two participants is established. First of all, each DRS representing an agent's, (i.e. hearer's or speaker's) cognitive state includes the two personal reference markers 'i' and 'you'. When 'i' is used in a DRS, it refers to the agent's self within that entire DRS. To refer to the other agent, 'you' is used. To make the link between speaker generation, linguistic content, and hearer recognition more explicit, presuppositions are marked by a presupposition label 'p_n', 'n' indicating a number. The labels increase the expressive power from an essentially first-order formalism to a higher-order formalism. Assertions are marked by 'a_n'. Similarly, DRSs inside the main speaker or hearer DRS are labeled 'drs_n'.

The reconciled version of DRT employs both Kamp's intention and beliefs spaces. However, the belief space is expanded to include the speaker's beliefs about the hearer's beliefs. The beliefs of an agent give the motivation for making an utterance, and the recognition of an utterance gives the hearer an insight into the agent's beliefs.

Another space or DRS is introduced to represent weaker belief, or '`acceptance`' space. This also includes the speaker's acceptance space as well as what the speaker takes the hearer to accept. Provided the speaker has sufficient information, the speaker can also have an embedded DRS within the acceptance space that represents what the hearer takes the speaker to accept. The same level of embedding is also introduced within the belief DRS when necessary.

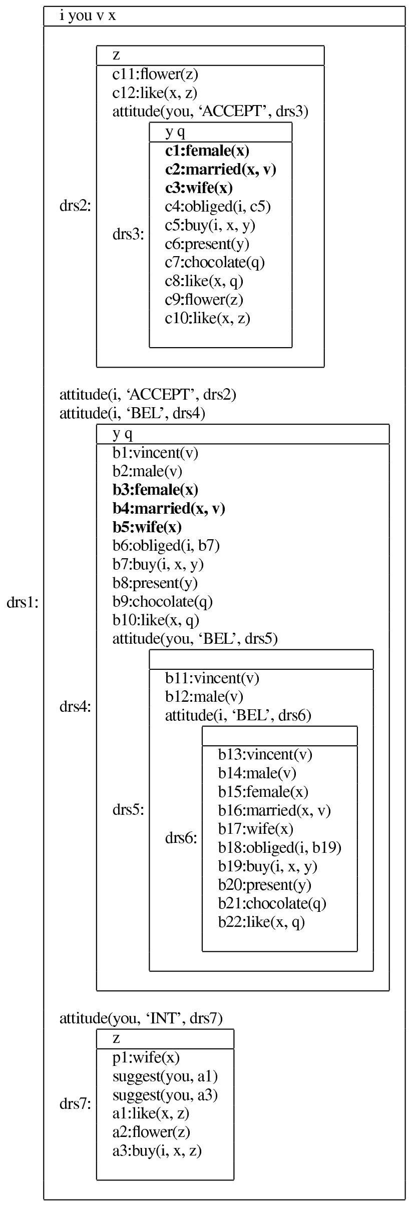


Figure 3: Speaker Recognition (After H2)

The intention space has been expanded to include the linguistic content provided by the current utterance, originally only represented in figure 1, to strengthen the link between an agent's intentions and the linguistic form uttered. The intention space also links the assertion with the presupposition that the particular assertion needs for linguistic realization and represents the dialogue act generated by making an assertion (Traum 1997). Believed information labelled 'b_n' inside a belief DRS or accepted information labelled 'c_n' inside an acceptance DRS can be either presupposed or asserted inside the intention DRS. As such, the labels in the intention DRS can only be 'p' or 'a'.

Intention space differs from the belief and acceptance spaces in that the intention space directly links to the current utterance being represented, whereas belief and acceptance spaces may include previous beliefs or accepted information. This gives the flexibility of being able to model information that the hearer has recognized but has not yet decided to accept or believe and, is therefore, not yet included in either the belief or acceptance space.

Figure 3 of speaker recognition in example 1 after H2 shows three embedded DRSs, acceptance DRS, drs2, belief DRS, drs4, and intention DRS, drs7. DRSs are referred to by the attitude describing them. For example, attitude(i,'BEL',drs4) refers to the DRS containing the speaker's beliefs, using the label for the belief DRS, drs4. The speaker's acceptance DRS, drs2, contains an embedded DRS for the hearer's acceptance space, drs3. Similarly, the belief space, drs4, contains space for the speaker's beliefs about the hearer's beliefs, drs5. The intention DRS, drs7, contains the recognized linguistic content of the utterance that the hearer made in H2.

Unlike example 2, the dialogue in example 1 has not provided the speaker with sufficient information to conclude that the hearer believes the assertions and presuppositions: the speaker has to buy a present, Vincent's wife likes chocolates, that the speaker will buy chocolates for Vincent's wife as a present, and there is such a person as Vincent's wife. After H2, the hearer has just suggested flowers as a present for Vincent's wife, which, given that the speaker has reason to believe the hearer is cooperating, leads the speaker to assume the hearer accepts the presupposition, Vincent has a wife. The propositions are thus represented in the the hearer's acceptance space, drs3, rather than his belief space, drs5. However, as the speaker initiated the topic of the conversation and indeed the type of present that Vincent's wife may prefer, the hearer has stronger grounds to believe that the speaker believes her utterances, drs6.

Example 2, on the other hand, where the hearer first questions the presupposition, Vincent has a wife, but later on agrees to buy a present for her, shows greater strength of belief attached to the presupposition which affects that commitment. By virtue of that commitment, the speaker can attribute greater strength to the hearer's beliefs about the presuppositions and assertions, which are represented in the speaker's beliefs about the hearer's belief, drs5, in figure 4 .

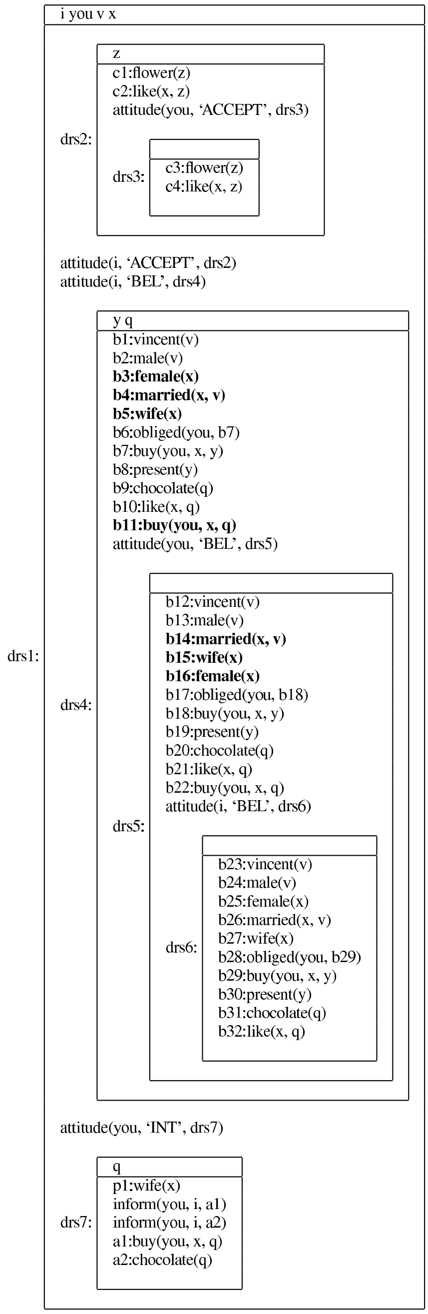


Figure 4: Speaker Recognition (After H3)

4 Conclusion

The paper attempted to represent the complex process of speakers recognizing utterances and using the linguistic information in forming mental representations of hearers' mental representations. This lead us to propose some modifications to DRT to offer compatible speaker/ hearer representations and handle examples where degrees of belief are needed.

References

Al-Raheb, Y. 2004. *Presupposition and Belief in DRT: Towards a New Implementation*. In: Ginzburg, J. and Vallduvi, E. Catalog 04: Proceedings of the 8th Workshop on the Semantics and Pragmatics of Dialogue, Barcelona, 144-145.

Kamp, H., van Genabith, J. and Reyle, U. 2005 *The Handbook of Logic*. Gabbay, D. and Guenthner, F. http://www.ims.uni-stuttgart.de/ hans/.

Traum, D. 1997. *Report on Multiparty Dialogue subgroup on Forward- looking Communicative Function*. Standards for Dialogue Coding in Natural Language Processing, Dagstuhl-Seminar Report no. 167.