Description Logics for Natural Language Processing

D. Fehrer, **U. Hustadt**, **M. Jaeger**, **A. Nonnengart**,
H. J. Ohlbach, **R. A. Schmidt**, **Ch. Weidenbach**, and **E. Weydert**
Max-Planck-Institut für Informatik,
Im Stadtwald
66123 Saarbrücken
Germany

1 Introduction

In this paper we focus on the application of description logics to natural language processing. In cooperation with the PRACMA Project[1] (SFB 314, Universität des Saarlandes, Germany) we have been developing a suitably extended knowledge representation system, called MOTEL.

In the late eighties inference in KL-ONE was shown to be undecidable. Since then the emphasis in research has been on developing and investigating systems that are computationally well behaved, i.e. are tractable or at least decidable.

As a result many commonly used description logics (also known as *terminological logics* or KL-ONE-*based knowledge representation formalisms*) have restricted expressiveness and are in their current form not suitable for natural language applications. This is evident, for example, from Schmidt [1993] who links knowledge representation with a relational approach to natural language semantics. For encoding knowledge formulated in a very limited fragment of English we already need the full expressive power of role constructs which have been eliminated in many languages.

In our approach to agent modelling and natural language processing we use an extension of the well-known description language $\mathcal{ALC}$. Our system MOTEL serves on one hand as a knowledge base for the natural language front-end, and on the other hand, it provides powerful *logical* representation and reasoning components. As our approach is logic based we hope that this enhances the overall capabilities of the natural language processing (NLP) system.

2 Natural Language Processing

The PRACMA project is concerned with pragmatic dialogue processing between two agents. These agents have the following properties:

(1) They communicate in natural language.

(2) They actively pursue complex goals, which may be conflicting.

(3) They have the means of analyzing (some of) the pragmatic content of what is being said, i.e., they have a deeper understanding of 'belief', 'intension' or 'argument'.

Figure 1 shows the architecture of the PRACMA system. The system is decomposed into modules. Each module is realized as an autonomous problem solver.

The module for recognizing propositional attitudes analyses certain linguistic expressions, e.g. modal verbs and modal adverbs. The results are stored in the agent model.

The module for assessment processing recognizes the positive and negative assessments of the agents towards certain objects, facts, and relations. The results are stored in the assessment knowledge base.

Instances of the plan processing module are the action planner controlling the agent's activities, e.g. collecting facts about objects, and the dialogue planner controlling the dialogue behaviour of the agent, e.g. opening the dialogue, raising a question. The planners rely on the agent model and the assessment knowledge base. In addition, they use the conceptual knowledge base, the argumentation strategy knowledge base, and the EGO knowledge base. The EGO stores behavioural patterns, e.g. the degree of cooperativity.

During the processing of a dialogue, each module can exist in multiple instantiations, called *actors*, working in parallel. The actors communicate and interact with each other using a protocol based on communication acts, i.e. on message exchange.

The test domain for the first prototype of the PRACMA system has been the processing of a dialogue between a car salesman and his customer. Figure 2 shows a small part of a dialogue and a schematic representation of its processing.

Note the following:

(1) The agents are not only exchanging facts, but beliefs, demands, etc. (e.g. 'So I don't *want* to buy it.').

(2) The beliefs of the salesman and the customer can contradict each other. The agents are able to detect such contradictions and can try to resolve them (e.g. 'No, that's not true.').

---

[1] PRACMA is short for 'PRocessing Arguments between Controversially Minded Agents.'

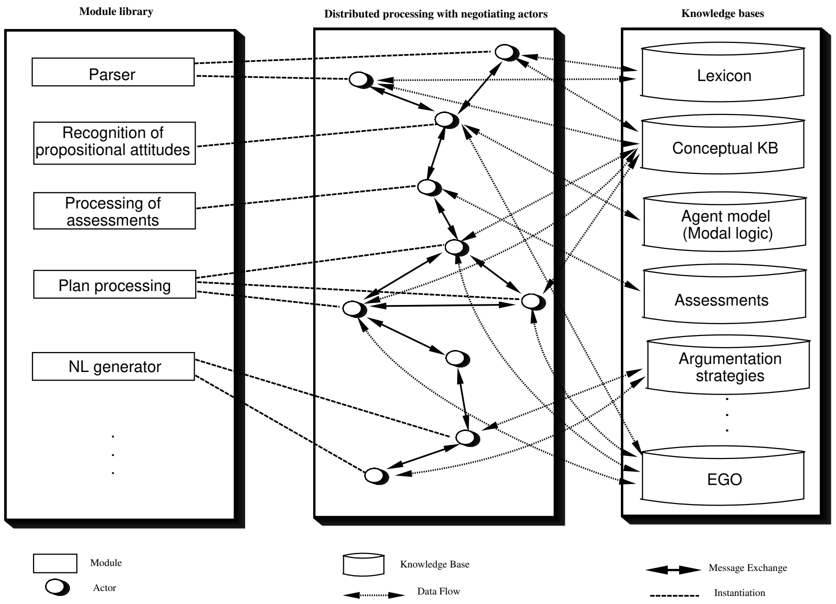

Figure 1: The architecture of the PRACMA system.

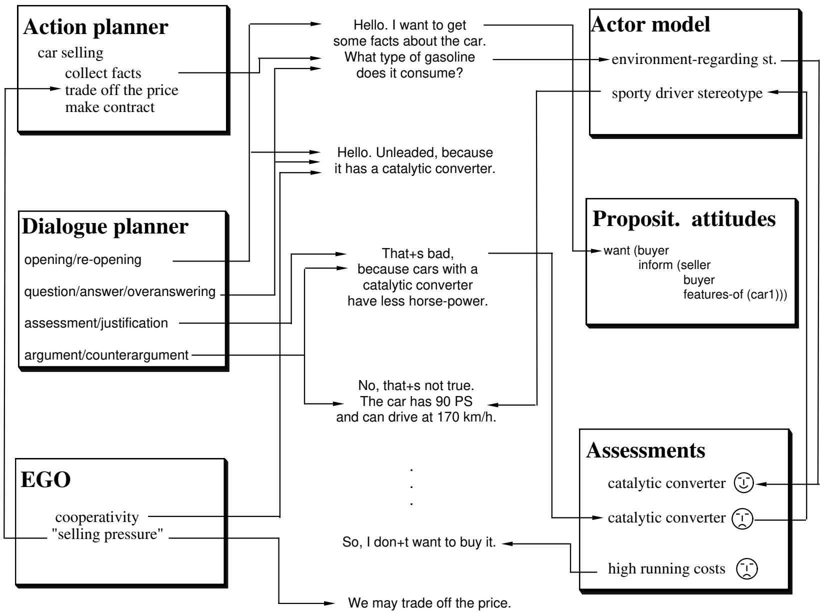

Figure 2: A sample dialogue.

(3) The agents are able to give reasons for their conclusions (e.g. 'Unleaded, because it has a catalytic converter.').

(4) Part of the information passed between the agents doesn't fit into a terminological system at all (e.g. 'That's bad.').

From the last item it is evident that terminological knowledge representation systems in their current form are not suitable for encoding the EGO and the assessment knowledge base. MOTEL is used to build the conceptual knowledge base and the agent model(s).

In the following sections we describe MOTEL and the different extensions we are working on.

3 Multi-Modalities

The traditional description logics can be used for representing common and individual knowledge about the world (domain of application). Recently description logics have been extended to allow the representation of the knowledge and the beliefs of multiple agents in one knowledge base [Donini *et al.*,1992; Kobsa,1992]. In MOTEL we formulate knowledge and belief as additional modal operators.

We are using $\mathcal{ALCNR}$ [Baader and Hollunder,1990] as a base language. That is, we assume three disjoint alphabets, the set of *concept names* C, the set of *role names* R, and the set of individual objects O. The set of *concept terms* (or just *concepts*) and *role terms* (or just *roles*) is inductively defined as follows. Every concept name is a concept term and every role name is a role term. Now assume that C, C_1, and C_2 are concepts, and R, R_1, and R_2 are roles. Then $C_1 \sqcap C_2$, $\neg C$, $\exists\, R.C$, $\exists_{\geq n}\, R.C$, and $\exists_{\leq n}\, R.C$ are concept terms, and $R_1 \sqcap R_2$, R^{-1}, $R|_C$ are role terms. The sentences of $\mathcal{ALCNR}$ are divided into *terminological sentences* and *assertional sentences*. If C_1 and C_2 are concepts and R_1 and R_2 are roles then $C_1 \sqsubseteq C_2$ and $R_1 \sqsubseteq R_2$ are terminological sentences. If C is a concept, R is a role, and O, O_1, and O_2 are individual objects then $O \in C$ and $(O_1, O_2) \in R$ are assertional sentences. As in [Baader and Hollunder,1990] we do not allow terminological cycles.

For the extended language Mod-$\mathcal{ALCNR}$ we assume in addition that we have an alphabet M of *modal operator names*. Also, there is a distinguished subset A of the individual objects, called the set of *agents*. We have a distinguished concept name '*all*' denoting the set of all agents with which we express mutual belief. The set of concepts and the set of roles of Mod-$\mathcal{ALCNR}$ contains all the concepts and roles of its sublanguage $\mathcal{ALCNR}$ and in addition it contains the concepts $\Box_{(m,a)}\, C$, $\Diamond_{(m,a)}\, C$, $\Box_{(m,C_1)}\, C_2$, and $\Diamond_{(m,C_1)}\, C_2$, and the roles $\Box_{(m,a)}\, R$, $\Diamond_{(m,a)}\, R$, $\Box_{(m,C)}\, R$, and $\Diamond_{(m,C)}\, R$, where m is a modal operator name and a is an agent name. The set of terminological and assertional sentences of Mod-$\mathcal{ALCNR}$ contains all the terminological and assertional sentences of $\mathcal{ALNCNR}$ and in addition it contains the expressions $\Box_{(m,a)}\, \Phi$, $\Diamond_{(m,a)}\, \Phi$, $\Box_{(m,C_1)}\, \Phi$, and $\Diamond_{(m,C_1)}\, \Phi$, where Φ is either a terminological or an assertional sentence.

We use a translational approach to provide the usual inference mechanisms, i.e. solving the consistency, the subsumption, the instantiation and the realization problem. Obviously, $\mathcal{ALCNR}$ knowledge bases can be translated into first-order logic theories. There are also well-known relational translation methods for modal logics. In [Hustadt and Nonnengart,1993] we have developed an improved translation method for Mod-$\mathcal{ALCNR}$ which provides an elegant translation of knowledge bases into first-order logic theories. In a prototypical implementation, the MOTEL system, we use a Prolog-based system with loop-checking as inference machine.

4 Quantitative Information

In MOTEL we use the cardinality-based approach proposed by Owsnicki-Klewe [1990] for dealing with number restrictions. Unfortunately, this approach is incomplete for languages in which concept disjointness is expressible.

The approach of Baader and Hollunder [1990], by contrast, provides a complete tableau method for $\mathcal{ALCNR}$, but has some disadvantages:

(1) The approach is not adequate for dealing with large numbers. Consider the following example: Suppose the universe consists of at most thirty objects. If there are at least twenty objects in C_1 and there are at least twenty objects in C_2, then there are at least ten objects in the intersection of C_1 and C_2.

The human ability to draw this conclusion is completely independent of the numbers we are using. Multiplying all numbers occurring in the example by a factor of 10 wouldn't make it any harder for us come up with the correct answer. Quite the opposite is true for the tableau method.

(2) The basic inference mechanism provided by tableau theorem provers is consistency checking for knowledge bases. This is adequate for answering queries that can be solved by checking the consistency of a suitably extended knowledge base, for example, for problems like subsumption, instantiation, and classification.

But the most suggestive class of queries for knowledge bases in $\mathcal{ALCNR}$, e.g. the question 'How many objects are in C_1 and C_2?' in the example above, cannot even be formulated.

A promising approach to quantitative reasoning with numerical quantifiers seems to be that of Hustadt et al. [1994], who investigate a translation technique which translates modal logics with graded modalities into a fragment of many-sorted first-order logic. For, $\mathcal{ALCNR}$ expressions can be associated directly with modal expressions.

5 Probabilistic Reasoning

Although Mod-$\mathcal{ALCNR}$ is a very sophisticated concept description language, the relationships among concepts that can be described are purely qualitative. Only inclusion, equality or disjointness relationships among concepts can be expressed. Jaeger [1994] investigates an extension of terminological knowledge representation languages that incorporates probabilistic statements. The language $\mathcal{PALC}$ based on $\mathcal{ALC}$ allows the following two additional kinds of sentences. *Probabilistic terminological sentences* are expressions $\mathrm{P}(C_1|C_2) = p$, where C_1 and C_2 are concept terms and $p \in [0,1]$. *Probabilistic assertional sentences* are expressions $\mathrm{P}(a \in C) = p$, where a is an element of O and $p \in [0,1]$. A knowledge base $\mathcal{KB}$ in $\mathcal{PALC}$ consists of a set $\mathcal{T}$ of terminological sentencesrestricted to $\mathcal{ALC}$, a set $\mathcal{PT}$ of probabilistic terminological sentences and a set $\mathcal{P}_a$ of probabilistic assertional sentences for every object name a: $\mathcal{KB} = \mathcal{T} \cup \mathcal{PT} \cup \bigcup\{\mathcal{P}_a | a \in \mathsf{O}\}$.

It is important to realize that these two kinds of probabilistic statements are completely different from each other. The former codifies *statistical information* that, in general, is obtained by observing a large number of individual objects and checking their membership of the various concepts. The latter expresses a *degree of uncertainty* of our belief in a specific proposition. Its value is usually justified only by a subjective assessment of likelihood.

Both kinds of probabilistic statements are interpreted in one common probability space which essentially consists of the set of concept terms that can be formed in the language of the given knowledge base. Defining all the probability measures on the same probability space allows us to compare the measure assigned to an object a with the generic measure defined by the given statistical information. The most reasonable assignment of a probability measure to a, we choose then, among all the measures consistent with the constraints known for a is the one that most closely resembles the generic measure. The key question to be answered, therefore, is how resemblance of probability measures should be measured. We chose the method of minimizing the cross entropy of the two measures.

6 Non-monotonic Reasoning

We have considered two different extension of the language $\mathcal{ALC}$ and its inference mechanisms to incorporate non-monotonic reasoning in MOTEL.

The first approach extends the language with an operator $\mathcal{A}$ of *assumability*. This operator can be applied to any concept term or role term, but it can only occur on the left-hand side of terminological sentences. The resulting language is called $\mathcal{ALCP}$.

A knowledge base $\mathcal{KB}$ in $\mathcal{ALCP}$ entails an assertion $a \in \mathcal{A}(C)$ iff $a \in \mathcal{A}(C)$ holds in all preferred models of $\mathcal{KB}$. Preference is defined with respect to the so-called *assumption order*.

In essence the implementation uses the negation as failure operator of PROLOG.

The second approach adds a new sentential operator T to $\mathcal{ALC}$ and a new subset declaration symbol $\sqsubseteq_{\mathsf{T}}$. If C_1 and C_2 are concept terms and Φ is a terminological sentence, then $C_1 \sqsubseteq_{\mathsf{T}} C_2$, and $\mathsf{T}\Phi$ are terminological sentences.

To provide a proof theory and a semantics for the extended language, we define a translation function mapping knowledge bases $\mathcal{KB}$ to default theories (W, D), where W is the set of first-order formulae and D is a set of supernormal defaults. The semantics of a knowledge base $\mathcal{KB}$ is the set of all possible extensions of (W, D). A knowledge base $\mathcal{KB}$ entails a sentence Φ iff Φ is entailed by every extension of (W, D).

7 Abductive Reasoning

Abduction was introduced by the philosopher Pierce as one of the three main forms of reasoning (the other two being deduction and induction). Abduction has widespread application in natural language processing systems. For example, Guessoum et al. [1993] describe the use of abduction for pronoun resolution. Most of the existing NLP systems use linguistic constraints for eliminating candidate referents, but it is widely recognized that non-linguistic knowledge is required to resolve ambiguities in general (c.f. the textbook example 'If the baby doesn't thrive on cows' milk, boil it'). More interesting for our testbed is the work of Quaresma and Lopes [1993] on abduction of plans and intentions in dialogues.

Hustadt [1993] proposes an abductive proof procedure for disjunctive logic programs with integrity constraints. Extending the class of normal logic programs to a class of programs including disjunction and integrity constraints permits arbitrary first-order problems to be stated in proper input format.

8 Reason Maintenance

There are at least two reasons why it is interesting to incorporate reason maintenance into a system like the one proposed here. The first is, that it may prove valuable not to dispose of the answers found to queries, but to keep them in order to be able to respond faster if the same queries or instances thereof occur again (similar to the use of lemmata in mathematics). As the knowledge base, however, is of dynamic nature, lemmata are only useful if their origins are remembered. The second reason is that we can't be sure that a knowledge base is globally consistent. So it is worthwhile looking for nogoods and reporting them, so that the master component is aware of them, or at least it can be guaranteed that in a single 'explanation' (proof) no inconsistent material is used (a kind of paraconsistency).

Fehrer [1993] shows how a reason maintenance system based on an arbitrary *basic logic* can be described logically. He also shows there, how an inference system can be obtained, given a calculus (axioms and set of inference rules) for the basic logic. As a special case we can get a system for $\mathcal{ALC}$.

At this stage this result is only of theoretical interest. The main advantage for using terminological logics, instead of full first order logics, lies in the fact that they have efficient algorithms for decidable fragments. The compound logic resulting from putting the reason maintenance onto $\mathcal{ALC}$ unfortunately cannot always make use of these algorithms (If we are content with only keeping track of the origins of lemmata generated so far, there is no problem, for the derived calculus inherits all the important properties from its ancestor, so the algorithms can be adapted in a simple manner). This is in essence due to the fact that in order to check for nogoods we have to generate *all possible derivations* of the falsum. If, however, we start with a possibly inconsistent knowledge base some *proof strategies* do not yield all possible derivations, for example, strategies incorporating set of support. But, since decidability as well as completeness is preserved in the compound system, it should be possible to devise algorithms with acceptable properties for that task.

9 Future Work

We want to focus on three parts of the architecture of natural language processing systems like PRACMA: The parser, the plan processing/NL generating modules, and knowledge representation system.

The natural language generating part of the system [Reithinger,1992] is a classical hierarchical planning system. In the current state, it doesn't make any use of the reason maintenance and abductive reasoning abilities of the knowledge representation system. The integration of these services of our system should improve the PRACMA system considerably.

The second prototype of PRACMA will use a parser translating natural language utterances into the semantic representation language $\mathcal{NLL}$. The language contains a first-order logic core, Boolean sentential operators, generalized quantifiers, plural reference expressions, λ-abstraction predicates, etc.

On the one hand, $\mathcal{NLL}$ provides more expressive power than we do in our terminological language. However, on the other hand, our language has syntactic constructs (like modal and probabilistic operators) not available in $\mathcal{NLL}$. If we would extend the syntax of both logics to a common language, we cannot provide correct and complete inferential mechanism for this logic. Therefore, we will have a core logic (based on Mod-$\mathcal{ALC}$) with correct and complete inference mechanisms and an extended logic (based on $\mathcal{NLL}$) with neither correct nor complete inference mechanisms.

10 Conclusion

We share the view of Doyle and Patil [1991] who argue for expressiveness as opposed to computational efficiency. Our experience with users interested in agent modelling and natural language simulations can be summarized as follows:

(1) Users want expressiveness.

(2) They want representation languages with more basic features than just concepts, roles and individuals (i.e. A-Box elements) and operations on these.

(3) And, they want special inference tools.

References

[Baader and Hollunder, 1990] F. Baader and B. Hollunder. $\mathcal{KRIS}$: Knowledge Representation and Inference System. System Description. Technical Memo DFKI-TM-90-03, Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, Germany, 1990.

[Donini *et al.*, 1992] Francesco M. Donini, M. Lenzerini, D. Nardi, A. Schaerf, and W. Nutt. Adding Epistemic Operators to Concept Languages. In B. Nebel, C. Rich, and W. Swartout, editors, *Proceedings of the KR '92*, pages 342–353, Cambridge, MA, 1992. Morgan Kaufmann.

[Doyle and S., 1991] J. Doyle and Patil R. S. Two theses of knowledge representation: language restrictions, taxionomic classification, and the utility of representation services. *Artifical Intelligence*, 48:261–297, 1991.

[Fehrer, 1993] Detlef Fehrer. A Unifying Framework for Reason Maintenance. In M. Clarke, R. Kruse, and S. Moral, editors, *Proceedings of the European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU '93)*, volume 747 of *LNCS*, pages 113–120, Granada, Spain, 1993. Springer.

[Guessoum *et al.*, 1993] A. Guessoum, B. Black, and J. Gallagher. Abduction for pronoun resolution. In P. Codognet, P. M. Dung, A. C. Kakas, and P. Mancarella, editors, *ICLP93 Postconference Workshop on Abductive Reasoning*, Budapest, Hungary, 1993.

[Hustadt and Nonnengart, 1993] U. Hustadt and Andreas Nonnengart. Modalities in knowledge representation. In C. Rowles, H. Liu, and N. Foo, editors, *Proceedings of the 6th Australian Joint Conference on Artificial Intelligence*, pages 249–254, Melbourne, Australia, 1993. World Scientific.

[Hustadt *et al.*, 1994] Ullrich Hustadt, H. J. Ohlbach, and Renate A. Schmidt. Qualified number restrictions and modal logic. Forthcoming MPI Report, 1994.

[Hustadt, 1993] Ullrich Hustadt. Abductive disjunctive logic programming. In P. Codognet, P. M. Dung, A. C. Kakas, and P. Mancarella, editors, *ICLP 93 Postconference Workshop on Abductive Reasoning*, Budapest, Hungary, 1993.

[Jaeger, 1994] Manfred Jaeger. Probabilistic reasoning in terminological logics. In J. Doyle, E. Sandewall, and P. Torasso, editors, *Proceedings of the KR '94*, Bonn, Germany, 1994. Morgan Kaufmann.

[Kobsa, 1992] Alfred Kobsa. Towards Inferences in BGP-MS: Combining Modal Logic and Partition Hierarchies for User Modeling (Preliminary Report). Bericht WIS-2, Projekt BGP-MS, Universität Konstanz Informationswissenschaft, June 1992.

[Owsnicki-Klewe, 1990] B. Owsnicki-Klewe. A cardinality-based approach to incomplete knowledge. In *Proceeding of the 9th European Conference on Artifical Intelligence*, pages 491–496, 1990.

[Quaresma and Lopes, 1993] P. Quaresma and J. G. Lopes. Abduction of plans and intentions in dialogues. In P. Codognet, P. M. Dung, A. C. Kakas, and P. Mancarella, editors, *ICLP93 Postconference Workshop on Abductive Reasoning*, Budapest, Hungary, 1993.

[Reithinger, 1992] Norbert Reithinger. *Eine parallele Architektur zur inkrementellen Generierung multimodaler Dialogbeiträge*, volume 1 of *DISKI - Dissertationen zur KI*. infix, St. Augustin, 1992.

[Schmidt, 1993] R. A. Schmidt. Terminological representation, natural language & relation algebra. In H. J. Ohlbach, editor, *Proceedings of the GWAI-92*, volume 671 of *LNAI*, pages 357–371, Bonn, Germany, 1993. Springer.