Bidirectional use of knowledge in the multi-modal NL access system XTRA

Jurgen Allgayer, Roman Jansen-Winkeln, Carola Reddig, Norbert Reithinger

SFB 314: AI - Knowledge-Based Systems, Dept. of Computer Science
University of Saarbriicken, D-6600 Saarbriicken 11, W.-Germany

{allgayer, jansen-winkeln, reddig, reithinger}@fblOvax.informatik.uni-saarland.dbp.de

Abstract

The acceptability and effectiveness of an expert system is critically dependent on its user interface. Natural language could be a well-suited communicative medium; however, current NL interfaces to expert systems show an impeding number of shortcomings. In this paper, it is shown how the XTRA system - a German NL interface between a user and one of various expert systems - employs a framework of novel AI techniques to overcome these shortcomings: it uses various highly interacting knowledge sources for domain-independent linguistic knowledge and for domain-specific world knowledge. This supports the view of NL as a universal communicative medium. In order to allow for a communicatively adequate dialog processing, knowledge sources are bidircctionaly used by both the analysis and the generation component. Both are enhanced with knowledge about the visual context of the dialog situation and about the capability to perform pointing gestures. SB-ONE - a knowledge representation formalism especially tailored for NL processing - deals with these types of knowledge. Together with some important tools of this formalism it constitutes the core of the system, whose performance is demonstrated with the help of an example dialog.

1 Theoretical Foundations of XTRA

Since expert systems - and other large knowledge based systems - are frequently employed in (prototypical) practical service, it turns out that the acceptability and, therefore, the effectiveness of an expert system (XPS) depends critically on its user interface ([Wahlster 86a,b, Hendrix 86, Tennant 861). Surely natural language would be a well-suited communicative medium if sufficiently mastered by the system. The XTRA system 1 is being developed in a basic research project which investigates open questions in the field of natural language processing, guided by the application as a (German) natural language (NL) interface to various expert systems [2].

Current NL interfaces show an impeding number of shortcomings. They are often tailored to the underlying system with respect to different aspects. One aspect is the direct translation of the input into the XPS specific knowledge representation ([Datskovsky Moerdler *et al.* 871), and vice versa in the case of generation. As a consequence, linguistic and domain-specific processing are intertwined. When the domain is changed, even linguistic processing has to be built up from the scratch ([Cohen and Jones 891). Another problem in NL processing is the lack of knowledge representation formalisms common and sufficient for all internal processes. One aim of the XTRA project is to overcome the mentioned drawbacks in NL interfaces to at least some extent. Special emphasis is laid upon the separation of linguistic and world knowledge, the bidirectionality of knowledge sources, and the combination of pointing gestures and written NL.

1.1 Linguistic and world knowledge

Knowledge about language is basically invariant, i.e. independent of a specific domain and, therefore, universal in any communicative situation. XTRA is equipped with two distinct (however: highly interacting) knowledge sources, the *Functional-Semantic Structure* (FSS) and the *Conceptual Knowledge Base* (CKB). The FSS's criteria for entailments and structuring are of linguistic nature, whereas the criterion of the CKB is its suitability for processing, e.g. drawing inferences and interacting with the XPS.

As [Wilensky 87] noted, it is rather pointless to claim that a sentence's representation based on pure linguistic knowledge is *equivalent* to its *meaning.* But it is beyond doubt (as is agreed by e.g. [Jacobs 85] and even [Wilensky 87]) that it has important merits in NL processing. We would by no means say that understanding or generating language is free from *extra-linguistic* influences; but it is necessary to explicitly represent the *linguistic* properties as well. In XTRA, the FSS is seen as the *intermediate structure* during analysis or generation of an utterance. Semantical well-formedness according to a Fillmore-likc deep case classification ordered in a taxonomic hierarchy and encoded in the knowledge representation language SB-ONE (see Fig. 1) is requested for *any* in/output. The construction of the CKB (sec Fig.2) should be independent of linguistic considerations to reflect the relations necessary for adequately modelling the domain of the underlying XPS.

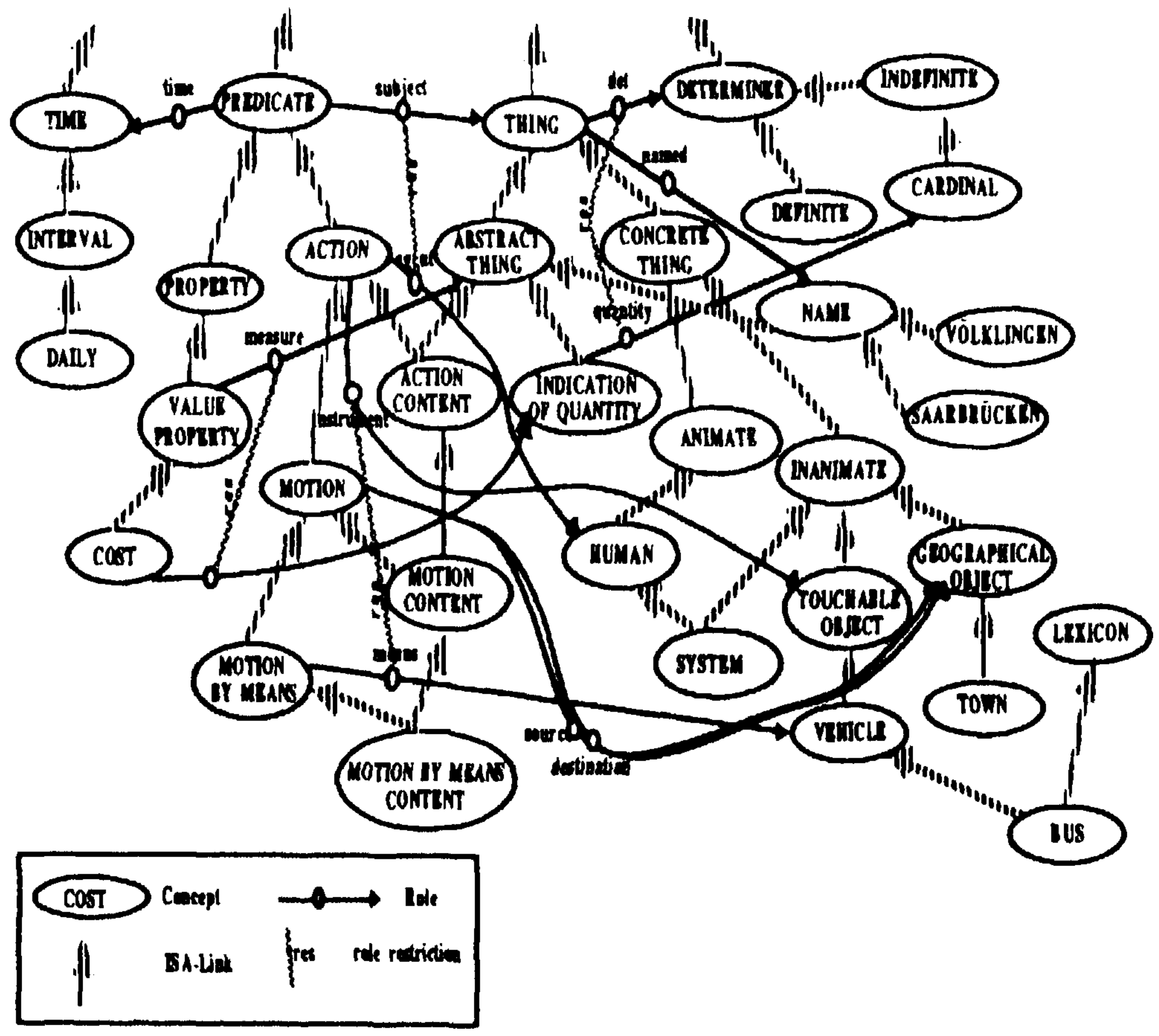


Figure 1: The taxonomy of the Functional-Semantic Structure (FSS) (fragment)

The FSS distinguishes between structures corresponding to *nouns* (concept *THING* with subconcepts in Fig.l) and those related to *verbs (PREDICATE).* Verb nominalizations combine both, corresponding to concepts within the *THING* hierarchy and being a subconcept of the respective *PREDICATE* as well. Consider, e.g., the concept *MOTION BY MEANS CONTENT* in the FSS, the nominalization of *MOTION BY MEANS*. However, within the CKB this distinction is no longer necessary, because they are mapped onto the same concept. For the above example it is *TO GO* or one of its subconcepts (see Fig.2).
Both the linguistic and extra-linguistic knowledge are strongly related to one another and both are encoded in SB-ONE. Translation rules map FSS onto CKB structures. They use SB-ONE structures as their preconditions and construct new SB-ONE individualizations as their result Tools of the representation system - especially matcher and realizer - support the rule interpreter, e.g. in constructing the conflict set for rule application.

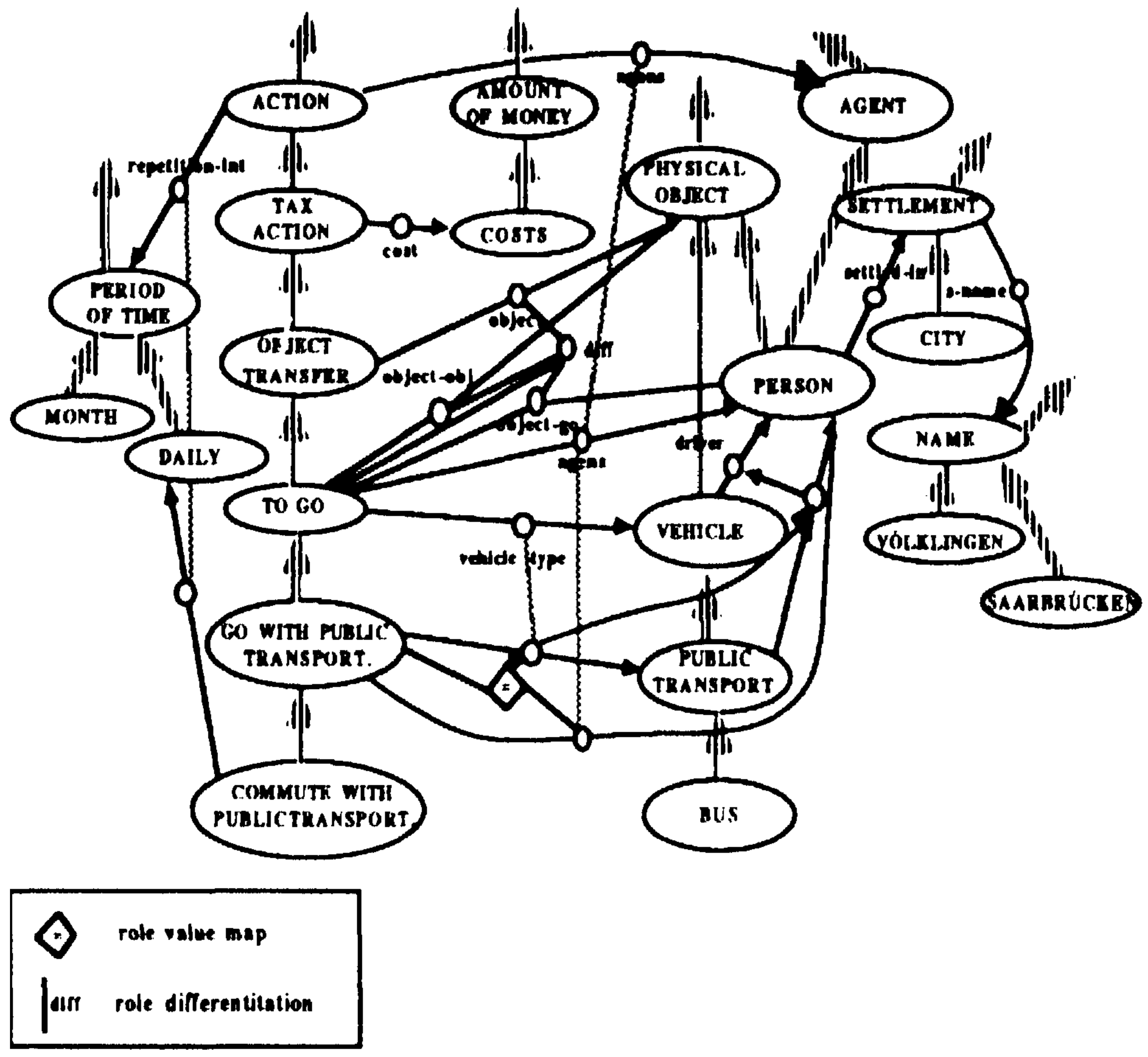


Figure 2: The Conceptual Knowledge Base (CKB)(fragment)

1.2 Bidirectionality of knowledge sources

Both analysis and generation share the Functional-Semantic Structure and the Conceptual Knowledge Base as their main knowledge sources. Sharing these sources facilitates generation and analysis of utterances with a high degree of interrelationships which is common in NL phenomena like mutual references or ellipsis. Due to the competence required for both analysis and generation, and the complexity necessary to cope with it, bidirectional use of knowledge bases is strongly recommended to avoid redundancy and inconsistencies (IShieber 88]). In addition to FSS and CKB, the dialog memory, the information about the visual context, and lexica arc shared. Although the strategy to generate a sentence is based on a parallel model ([Reithinger 88]) and differs from the sequential approach of the analysis ([Allgayer and Reddig 86]), many core algorithms were kept independant: morphological processing ([Finkler and Neumann 88]), interpretation of pointing gestures ([Allgayer 86]), the unification-based syntax formalism PATR ([Shicber *et al.* 83], [Finkler and Neumann 89b]), and the multi-modal front-end ([Allgayer *et al.* 89]).

1.3 Combining linguistic expressions and pointing gestures

Although NL fits the communicative needs of user and system in most situations, additional possibilities in specifying objects may simplify the interaction especially whenever a shared visual context is part of the dialog situation. As the visual context, XTRA offers a domain-specific graphics on the terminal screen to which both the user and the system can refer to by pointing gestures. In a current application of XTRA - NL access system to a tax advisory XPS - a tax form is displayed to parts of which the user as well as the system can refer to, especially with pointing gestures, and fill in data being directly delivered to or provided by the XPS. On the linguistic level, pointing gestures accompany noun phrases or adverbial phrases. The gestures are to be analyzed and generated as modifiers of the respective structures of the Functional-Semantic Structure (sec Fig.3).

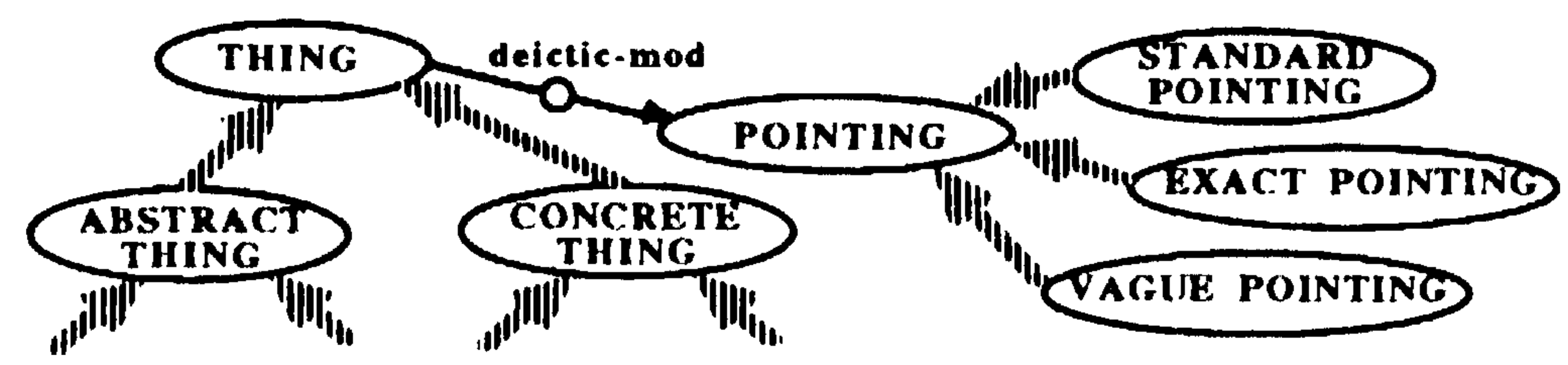


Fig.3: Representation of pointing gestures in the FSS_{TBox}

They are processed using a specific knowledge source - the form hierarchy (FH) - which contains structural information about the visual context. In order to understand what is to be seen on the screen in terms of world knowledge, the elements of the FH are connected to the CKB through a framework of links. This framework relates those regions of the form in which strings or numbers may be filled in to individualized concepts of the CKB. The pre-defined parts of the form are equipped with a link to general concepts and may get an additional link to specific, individualized CKB structures after specifying concrete data for these regions (see Fig.6). This corresponds to a moderate view of representing visual objects which takes on the "rigid dcsignator"-function only for names and numbers and performs a "word-to-meaning"-correspondence for strings visible on the form.

1.4 The knowledge representation language SB-ONE

We developed our own representation formalism for XTRA's knowledge called SB-ONE ([Kobsa 89, Profitlich 89]). It is a derivate of KL-ONE which is widely used in the field of NL systems. Along with other KL-ONE languages, it shares some features which distinguish them from frame-like formalisms usually applied in XPSs: number restrictions and role value maps which are of special importance in NL processing. In addition to KL-ONE, SB-ONE provides default value restrictions, optional roles, and a context mechanism which together allow for adequate formulation of linguistic as well as world knowledge. The context mechanism, which at the moment divides sets of T- and A-Box structures into separate subnets - so-called "partitions" - is being developed to implement a user modelling component ([Kobsa 881).
Furthermore, specific SB-ONE tools support the tasks occuring in NL processing. Besides classifier and realizer, the matcher is one of them which constitutes a part of SB-ONE's ASK component It maps pattern structures - containing individualized concepts and roles enhanced by so-called variable concept and roles - onto SB-ONE individualizations. The matcher is used within the FSS-to/from-CKB translation and for referent identification purposes. The matcher returns a list of those structures that are agreeable with the pattern containing the actual description (for further explanation sec [Aue *et al.* 89J).

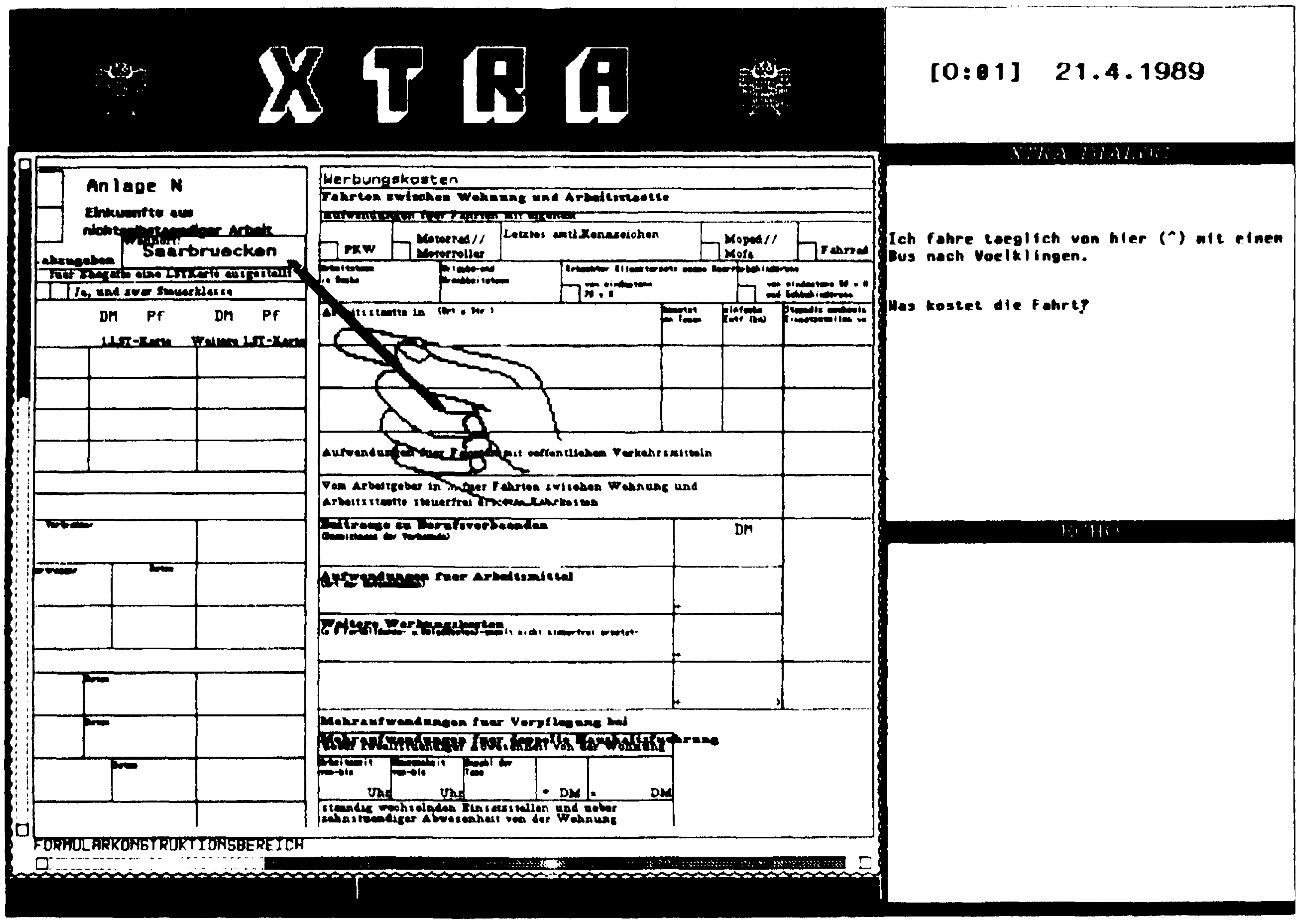


Figure 4: Screen with the example dialog (fragment)

2 XTRA in Action - an Example Dialog

XTRA can roughly be separated into three main tasks: the analysis of the user's input, the evaluation of the meaning representation which comprises inferences and the interaction with the expert system, and the generation of NL dialog contributions. We illustrate these steps by an example dialog.
In its current application, XTRA supports the user's filling out a (German) annual tax withholding adjustment form ([Beiche 87]). Figure 4 shows (a part of) the front-end with the visual context and the user's pointing gesture accompanying the first sentence and the windows for NL in/output.

Example Dialog: U: "Ich fahre täglich von hier (^) mit einem Bus nach Völklingen."
("I commute daily from here (^) Völklingen by bus.")
S: "Was kostet die Fahrt?"
("How much is the trip"?)
U: "Sie kostet 35DM im Monat."
("It is 35DM per month.")
S: (computes and fills in the yearly costs at the appropriate position in the form)

The pointing gesture to be seen in Figure 4 is marked syntactically by (^).

2.1 Building up the meaning representation

The morphological analysis is performed by MORPHIX, which is integrated into the unification-based parser SB-PATR. It uses a constituent structure grammar for syntactic analysis and derives Directed Acyclic Graphs (DAGs). One of the two alternative DAGs contains the following information about the phrase with the pointing gesture:

```
POBJ: [ADVERB: [DEICTIC: [POS: 6
                          STEM: (^)]]
               POS: 5
               SEMCLASS: LOC
               STEM: HIER
               TYPE: DEICTIC]
PREP: [CASERESTR: 3
       POS: 4
       STEM: VON]
SEMCLASS:<0 FSET STRUCTURE POBJ ADVERB
          SEMCLASS>
```

Starting from the syntactic results, the sentential-semantic analysis constructs the corresponding individualized part of the Functional-Semantic Structure (see Figure 5). Therefore it uses the semantical lexicon, which is an extension of the FSS and introduces domain-dependent word-concepts using the domain-independent terminology provided the FSS_{TBox} (see Fig.1, concepts ***LEXICON*** and ***BUS***). It contains information of both the linguistic input as well as the pointing gesture. We would like to stress that this structure will be built up identically for every occurence of the respective sentence in any domain.

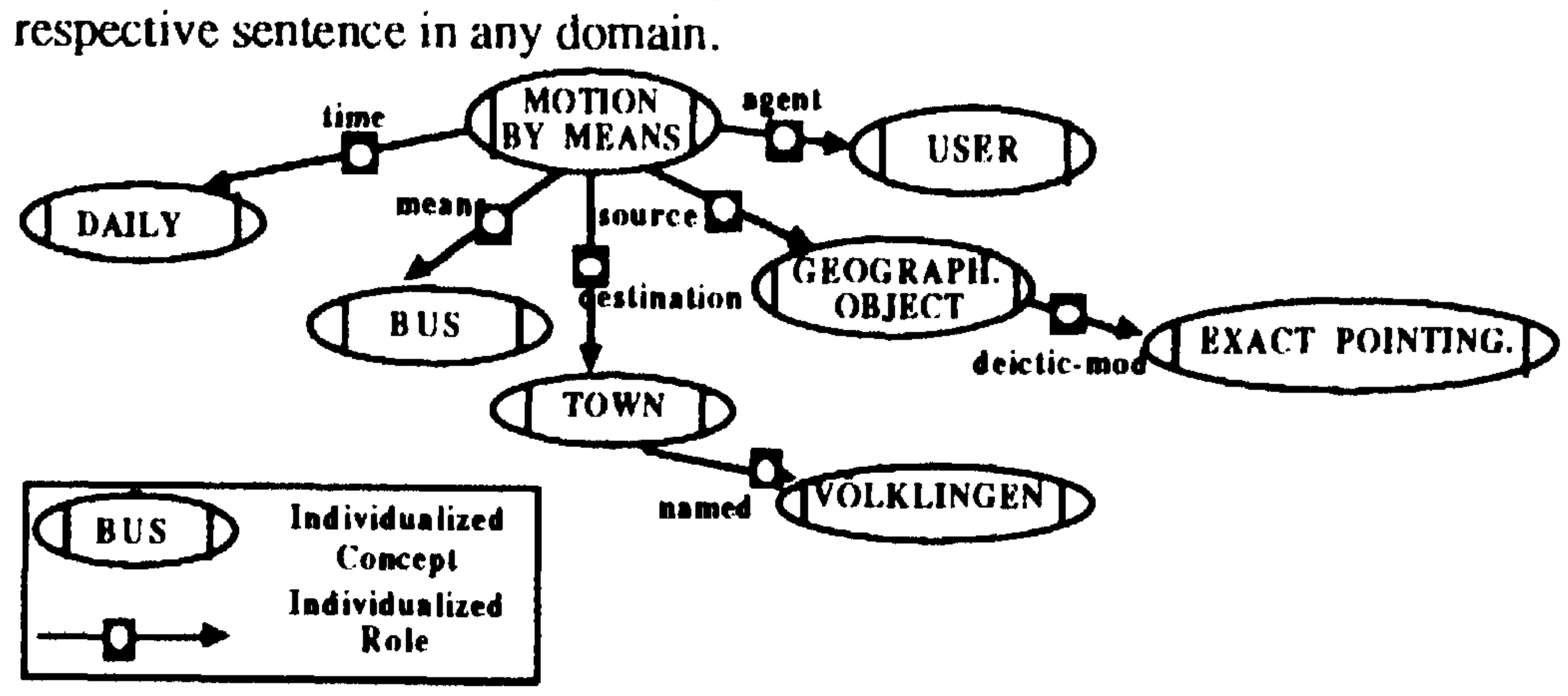
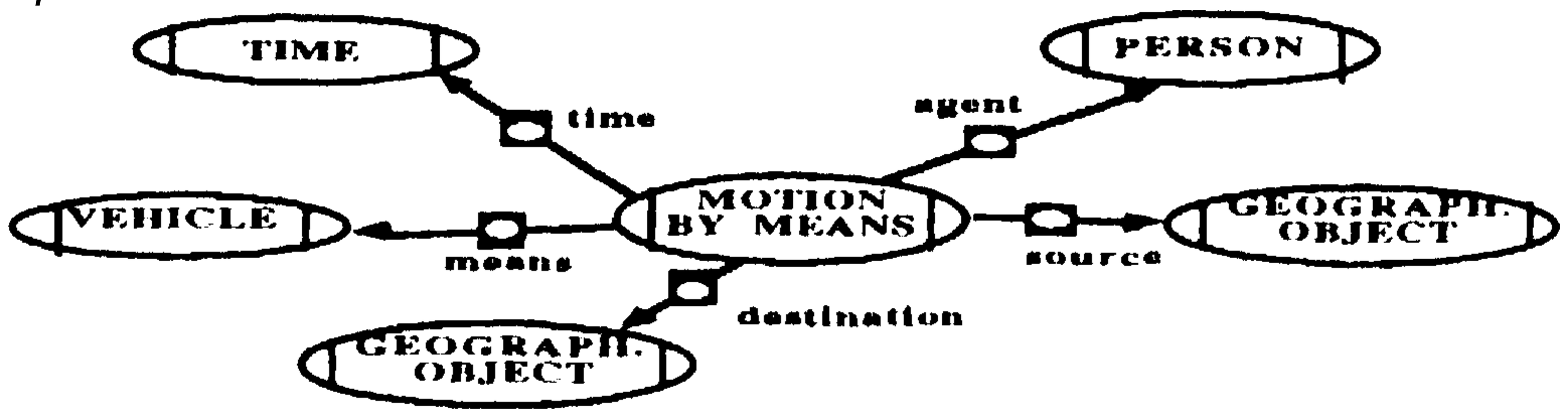


Fig.5: Resulting FSS_{ABox}

The next step is the rule-based translation process between the domain independent, language oriented structure FSS and the domain specific CKB. Here is one of the rules needed for this process:

.-precondition




```
:action '(TO GO
        (object-go              (<- agent))
        (destination            (<- destination))
        (source                 (<- source))
        (repetition-interval    (<- time))
        (vehicle-type           (<- means))
```

The precondition refers to a FSS pattern partition. It is matched successfully against the individual FSS (called FSS_{ABOX}). The action part consists of the reference to a CKB concept *(TO GO)* which is to be individualized as the translation of the FSS concept *MOTION BY MEANS.* For the roles of this individualization, the fillers arc specified in terms of translations of the corresponding FSS roles. For example, the tiller of the role *object-go* at *TO GO* has to be the result of the translation (indicated by <-) of the FSS role *agent* at *MOTION BY MEANS.* This translation process results in individualized CKB structures which are stored in the user-says partitions. The realizer is applied to this structures in order to attain their most specific meaning in terms of world knowledge.

The next task is the referent identification. The necessary information about searching for an existing referent or creating a new one has already been annotated in the user-says partitions during the translation process. These annotations reflect features in the NL input such as definitencss. Together with possible referents delivered by the dialog model, they guide the matching process of user-says against CKB_{ABOX}. The resulting referent candidates are filtered and ordered according to various criteria, e.g. best-fitting matching result, or referents which had been mentioned last recently. The remaining structures arc used to extend the CKB_{ABOX} in order to get a unique and consistent meaning representation of the input.

2.2 A special look at deictic analysis

The input was accompanied by a pointing gesture. Therefore, FSS_{ABox} carries information about the location and granularity of this gesture. During the referent identification process, TACTILUS ((Allgaycr 87]) analyzes this information together with the CKB structure of user-says, which is the translation of the phrase "von hicr (^{A}) ". Starting from the CKB structure, the corresponding parts of the form arc extracted by exploiting the links between the form hierarchy and the conceptual knowledge base. All elements contained in the exploited section of the form hierarchy are rated with evidences for being pointed on, based on the granularity and location of the gesture. Following the links back into the CKB, we Find those CKB structures which will be passed to the referent identification process as being an ordered list of possible referents. For the example input, the unique candidate is *SAARBROCKEN* as the filler of the *name* role for *CITY-2.*

Figure 6 shows the relevant sections of the CKB_{TBOX} (upper part), the meaning representation of the input (lower part), and the form hierarchy, linked together. A *D'* marks roles with default Fillers.

2.3 Evaluating the meaning representation

The system has successfully analyzed the input sentence. The aim now is to provide the XPS with data to be extracted from the meaning representation. Therefore, several knowledge sources arc used, including knowledge about the structure of the XPS's knowledge base and its connections to the CKB. In our case, a successful transfer requires further individualization of the neccessary role *costs* for which - at the moment - only a default filler is present. This constellation justifies the marking of the *costs* role (including its filler) to be asked for.

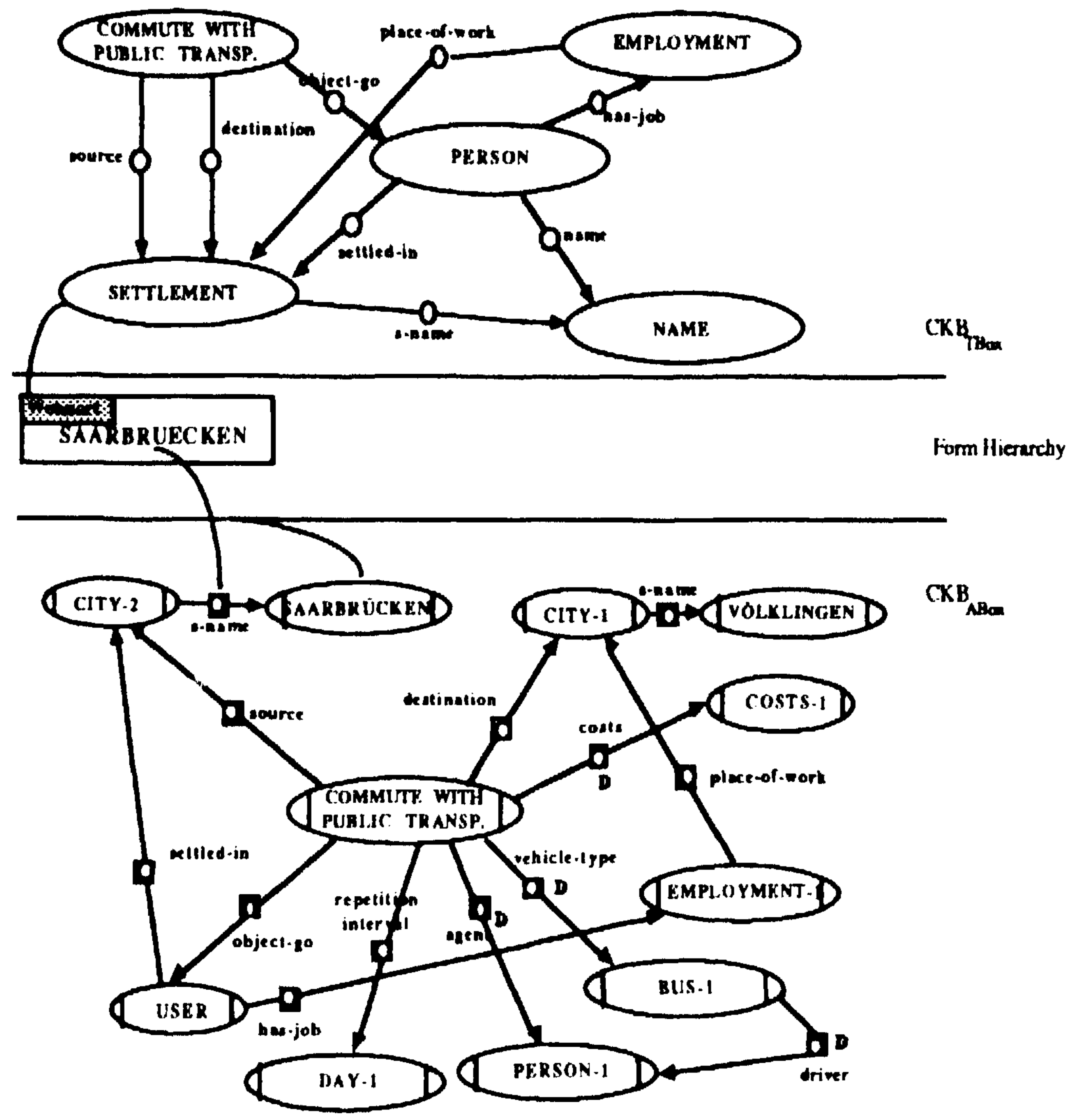


Figure 6: The meaning representation of the first sentence

2.4 Verbalizing the system's reaction

POPEL [Rcithinger 88J, the generator of the XTRA system, consists of two parts, the what-to-say component and the how-to-say component. The first decides on the basis of marked CKB structures which parts of them must be included in the output. Since only complete SB-ONE structures can be verbalized, it additionally selects the dominating concept of the role *costs (COMMUTE WITH PUBLIC TRANSPORTATION)* Furthermore, the markers determine the type of illocutionary act to be choosen - a question in our case.

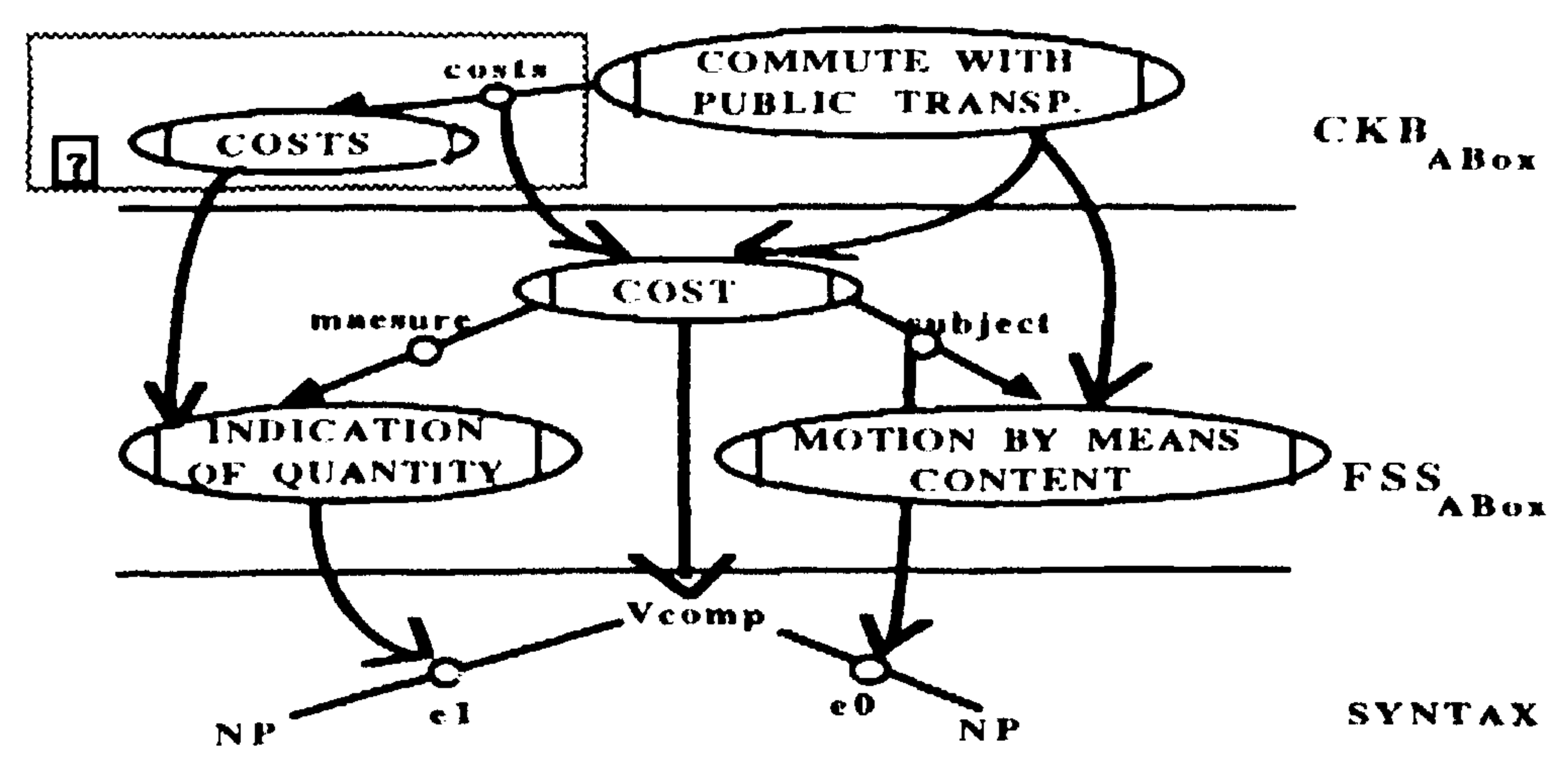


Figure 7: The CKBABox to be verbalized by the generator with corresponding FSSABox and dependency structure

The selected CKB individualizations are passed incrementally to the how-to-say component, which constructs the surface structure in parallel. Again, an individualized FSS structure represents intermediate results, built up by rules corresponding to those used during analysis. A remarkable aspect is the nominalization of *COMMUTE WITH PUBLIC TRANSPORTATION,* mapped onto *MOTION BY MEANS CONTENT* instead of *MOTION BY MEANS.* This is due to the presence of the predication *COST.* In the next step, the FSS structure is transformed into a dependency based syntactic structure using a unification-based grammar (realized in a parallel extension of PATR IFinkler and Neumann 89b]). Finally, the resulting preterminal chain is inflected with MORPHIX.

2.5 Activating the XPS

As an answer to XTRA's request about the travel expenses, the user types in: "Sie kostet 35DM im Monat." (It *is 35DM per month")* The analyzer operates analogously to the first sentence. In the CKB structure, *35* is individualized as the filler of the role *number* from *AMOUNT OF MONEY* which, in turn, becomes the role filler of the already existing role *costs* of the CKB_{ABOX}. Now, the transfer of data to the XPS is possible: the yearly costs required by the XPS are computed, passed over and inserted at the appropriate position in the form.

3 Future Work

Up to now, a prototype of XTRA has been implemented in Common Lisp on a Symbolics 3640. This will serve as basis for further development of the system. Besides improving already existing modules, e.g. the context inheritance mechanisms for SB-ONE, we will concentrate on the following topics.

A knowledge aquisition component will be developed in order to facilitate extensions within the existing knowledge sources and the adaptation to new XPSs. With the use of examples, this component allows for an integrated access to and exploitation of XTRA's multiple knowledge sources during the system's NL interaction with the user (JansEn-Winkeln 88]). On top of SB-ONE, a logic-oriented TELL and ASK interface will be developed which allows for the representation of NL quantifiers, disjunctions, and negative information as well as for inferences to be drawn from it (IAllgayer and Reddig 89]). Just as with the analysis of the input, the generation component will be equipped with the gesture generator ZORA ([Jung *et al.* 89]), which enables the use of pointing gestures by the system as well. Again, the bidirectional use of data and algorithms of TACTILUS supports the development of the new component. A module for dialog structuring and processing will extend the linguistic capabilities of the system.

This research has been supported by the German Science Foundation in its Special Collaborative Research Programme on Artificial Intelligence and Knowledge Based Systems (SFB 314).

[1] XTRA stands for "eXpert TRAnslator".

[2] Currently, XTRA has access to an expert system in the tax domain; the next will be a classification system for mushrooms.

References

[Allgayer, 86] Allgayer, J: "Eine Graphikkomponente zur Integration von Zeigehandlungen in naturlichsprachliche KI-Systeme." In: Proc. der 16. GI-Jahrestagung, Berlin etc., Springer, 1986.

IAllgayer and Reddig. 86] Allgayer, J. and C. Reddig (1986): "Processing Descriptions containing Words and Gestures - A System Architecture" In: C.-R. Rollinger and W. Horn, eds.: GWAI-86 and 2. Osterreichische Artificial Intelligence Tagung. Berlin etc., Springer, 1986.

[Allgayer and Reddig, 89] Allgayer, J. and C. Reddig (1989): "What KL-ONE Lookalikes need to Cope with Natural Language." Proc. of the Workshop on Order-Sorted Representation Formalisms, Springer, Berlin, to appear.

[Allgayer et al., 89] Allgayer, J., K. Harbusch, A. Kobsa, C. Reddig, N. Reithinger, and D. Schmauks: "XTRA: A Natural-Language Access System to Expert Systems." In: International Journal of Man-Machine Studies. To appear.

[Aue et al., 89] Aue, D„ S. Heib and A. Ndiaye: "SB-ONE Matcher: Systembeschrcibung und Benutzeranleitung." Memo Nr. 32, University des Saarlandes, FB Informatik, D-6600 Saarbriicken, W-Germany, 1989.

[Beiche, 87] Beichc, H.-P.: "Lst-1: Ein wissensbasiertes System zur Durchfuhrung und Bcrechnung des Lohnsteuerjahrcsausgleichs." In: E. Buchbergcr and J. Retti, eds.: 3. Osterreichische Artificial-Intelligence Tagung. Berlin etc., Springer, 1987.

ICohen, 88] Cohen, R. and M. Jones: "Incorporating User Models into Expert Systems for Educational Diagnosis." In: A. Kobsa, W. Wahlstcr, eds.: User Models in Dialog Systems. Berlin etc., Springer, 1988.

[Datskovsky Moerdler et al., 87) Datskovsky Moerdler, G, K.R. McKeown and J.R. Ensor: "Building Natural Language Interlaces for Rule-Based Expert Systems." In: Proc. 10th IJCAI, Milan, Italy, 1987.

[Finkler and Neumann, 88] Finkler, W. and G. Neumann: "MORPHIX: A Fast Realization of a Classification-Based Approach to Morphology." In: H. Trost, ed.: 4. Osterreichische Artificial-Intelligence Tagung, Wien, Austria, 1988.

[Finkler and Neumann, 89a] Finkler, W. and G. Neumann: "SB-PATR: Systembeschreibung und Bcnutzeranleitung." Memo, Universitat des Saarlandes, FB Informatik, D-6600 Saarbriicken, W-Germany, 1989.

[Finkler and Neumann, 89bJ Finkler, W. and G. Neumann: "POPEL-HOW - A Distributed Parallel Model for Incremental Natural Language Production with Feedback." Proc. of 11th IJCAI, Detroit, MI, 1989.

[Hendrix, 86] Hendrix, G.G.: "Q&A: Already a Success?" In: 11th COLING, Bonn, West Germany, 1986.

[Jacobs, 85] Jacobs, P. S.: "PHRED: a generator for natural language interfaces." Computational Linguistics, 11 (4), 1985.

[Jansen-Winkeln, 88] Jansen-Winkeln, R.M.: "WASTL: An Approach to Knowledge Aquisition in the NL Domain." In: EKAW'88, Bonn, W-Germany, 1988.

[Jung et al., 89] Jung, J., A. Kresse, N. Reithinger, R. Schafer: "Das System ZORA - Wissensbasierte Generierung von Zeigegesten" Submitted to GWAI 89.

[Kobsa, 88] Kobsa, A.: "User Models and Discourse Models: United they stand..." Computational Linguistics, 14(3), 1988.

IKobsa, 88] Kobsa, A.: "The SB-ONE Knowledge Representation Workbench." In. J. Sowa, ed.: Formal Aspects of Semantic Networks, Los Altos, CA, Morgan Kaufmann, 1989.

[Profitlich, 89] Profitlich, H.J.: "Das SB-ONE Handbuch", Memo Nr. 33, Univcrsitat des Saarlandes, FB Informatik, D-6600 Saarbriicken, W-Germany, 1989.

[Reithinger, 89] Reithinger, N.: "POPEL - A Parallel and Incremental Natural Language Generation System." Paper presented at the 4th Int. Natural Language Generation Workshop. Internal Working Paper Nr.2, Universitat des Saarlandes, FB Informatik, D-6600 Saarbriicken, W-Germany, 1989.

IShieber et al., 83] Shieber, S.M., H. Uszkoreit, F. Pereira, J. Robinson and M. Tysun: "The Formalism and Implementation of PATR II." In: B. Grosz and M. Stickel, eds: Research on Interactive Acquisition and Use of Knowledge, SRI International, Stanford, CA, 1983.

[Shieber, 88] Shieber, S.M.: "A Uniform Architecture for Parsing and Generation". In: Proc. 12th COLING, Budapest, Hungaria, 1988.

[Tennant, 86] Tennant, H.: "The Commercial Application of Natural Language Interfaces." In: Proc. llth COLING, Bonn, West Germany, 1986.

[Wahlster, 86a] Wahlster, W.: "Natural Language Interfaces: Ready for Commercial Success?" Panel Introduction, Proc. 11thCOLING, Bonn, West Germany, 1986.

[Wahlster, 86b] Wahlster, W.: "The Role of Natural Language in Advanced Knowledge-Based Systems." In: H. Winter, ed.: Artificial Intelligence and Man-Machine Systems. Berlin etc., Springer, 1986.

[Wilensky, 87] Wilensky, R.: "A Framework for a Theory of Meaning." TINLAP-3, New Mexico State University, Las Cruces, NM, 1987.