

GENERATING REFERRING EXPRESSIONS USING MULTIPLE KNOWLEDGE SOURCES

Russell Block
Universität Hamburg
Zentrales Fremdsprachen-
Institut
Von-Melle-Park 5
2000 Hamburg 13
F.R.G.

Helmut Horacek
Universität Bielefeld
Fakultät für Linguistik
und Literaturwissenschaft
Postfach 8640
4800 Bielefeld 1
F.R.G.

Abstract

In this paper we present a brief look at some of the knowledge-based processes used in generating referring expressions in the natural language advisory system WISBER. Although WISBER is fully capable of exploiting syntactic information to generate contextually appropriate references, the work described here concentrates on the use of conceptual and contingent knowledge about objects in the domain of discourse to generate natural-sounding references. A short description of the knowledge sources available is followed by examples of the processes that transform "deep structures" encoding system intentions into verbalizable form. Finally, we discuss a number of problems of specifier selection and their solution within a knowledge-based framework.

1. Introduction

Competent handling of referring expressions is an important prerequisite for skillful natural language (NL) processing. Not surprisingly, considerable effort has been invested in resolving references, but comparably little attention has been paid to the generation of referring expressions as yet. Moreover, approaches in generation have concentrated mainly on syntactic problems and pronominalization issues.

Our contribution, however, lies in putting significantly more emphasis on semantic and pragmatic aspects, and in generating (ordinary) anaphoric noun phrases in addition to pronouns. We achieve this by updating and exploiting the content of the knowledge sources

of the advisory system WISBER [12], which our generator is part of. WISBER is a fully implemented German NL system covering the whole spectrum of NL processing. Its domain of application is financial investment.

An outline of the knowledge sources and the coordination of the associated processes has been given in [6]. In this paper we will concentrate on the motivation of criteria for choices in generation and on the presentation of some (in part) tricky types of reference generation which our system is able to master. To start with, we will characterize the knowledge sources in our system and then describe the subprocesses in the generation module involved in the creation of referring expressions.

2. Knowledge sources involved

Conceptual knowledge is expressed in a T-Box [4] containing structured terminological knowledge about the world, and by an A-Box [19], which contains assertions about entities referred to in the dialog (both are in the tradition of hybrid KL-ONE based knowledge representation systems [7]). The partner model [20], for instance, is realized as a particular context in the A-Box (which contains the beliefs of the system about the propositional attitudes of the user). Additionally, dialog specific (heuristic) knowledge is expressed in terms of derivational rules for inferring additional propositional attitudes of the agents involved [10]. All these components share the same ontology which is object-oriented and particularly well-suited for representing conceptual knowledge and for making inferences. Consequently, there are significant structural differences in comparison to lexically-based representations. The prin-

ciples on which the design of the ontology is based are outlined in [14].

The bridge between the conceptual and the lexical levels is established by a 'semantic' lexicon [13], which contains entries for each word meaning (and for grammatical functions) which comprise the conceptual entity primarily addressed, a schema type which indicates how the associated conceptual structure has to be built and a few parameters which constrain the lexical and the conceptual environments. As with the syntactic-semantic lexicon used in the system VIE-LANG [21], the entries can be interpreted bidirectionally and the resulting partial structures are composable. The major difference, however, lies in the concentration on a limited set of schema types in our approach.

Additionally, an elaborate feature system [5] has been developed, composed of grammatical and semantic features for defining objects. The dialog memory [5] contains objects defined by these features, thus providing a link between objects from the world of discourse (the SEMs) and the natural language expressions (names) used to refer to these objects (the REFOs). Quite naturally, one SEM may be expressed by several REFOs in the course of the conversation, but we will see that things may occasionally be more complicated.

3. Associated processes

We restrict our presentation of the generation process to the phases starting with an (initial) representation of the utterance to be produced (expressed in IRS (*Interne Repräsentations-Sprache*) [3], WISBER's dedicated language for expressing utterances on the semantic-pragmatic level) up to a level comparable to functional descriptions (which is called IRS-F [5]). The processes involved are transformation of IRS-expressions on a purely terminological level (by the component called FTRANSLATE [2]), the selection of appropriate descriptions for entities (which works similarly to the NP-generation component in HAM-ANS [17]) and the transformation from the conceptual level to the lexical level (by the verbalization component also referred to in [13]).

FTRANSLATE makes it possible to replace appearances of a special concept by a more general one (augmented by additional descriptions to maintain terminological equivalence). In addition, a role associated with complex meaning can be re-expressed by a construct consisting of other roles and associated concepts. The terminological equivalence is defined by means of a structural description, an

element in the T-Box language. Hence, entirely new elements might be included in the specification of an utterance generated this way.

In the NP-generation component, the descriptions of (semantic) objects are expanded with additional properties so that the extended descriptions are uniquely identifiable (can clearly be distinguished from other objects found in the dialog memory with which they could potentially be confused if the objects were considered out of context, [6]). As in [17, 18] the observations that people have preferences when using properties to characterize objects are taken into account (e.g., they prefer color over size over age). Notice that, in contrast to the other approaches, these descriptions are not necessarily reflected entirely in the corresponding surface form in our system. It is up to the verbalization process to select an adequate realization (e.g., a paraphrase), taking the overall dialog context into account. This refers, in particular, to pronominalization decisions. Hence, our NP-generation component always produces a description consisting of a class name and a set of uniquely identifying properties.

The transition between the conceptual and the lexical levels, which we have termed verbalization, is described by a small set of schemata which serve to bridge differences in granularity (ZOOM schemata) and in the degree of explicitness (SUBSTITUTE schemata). Hence, this transition may involve considerable restructuring. In particular, partial mappings of conceptual structures may be collapsed on the lexical level (one of them may be *substituted* for another). Moreover, some parts of the (full) specification may be left un verbalized because the overall context indicates that the reduced message is comprehensible without any loss of information (i.e., the superfluous parts can be assumed to be contextually recoverable). The extra effort involved in creating these expansions in the NP-generation component is necessary to keep the flow of control between the processes simple (in fact they are sequentially ordered).

The verbalization process and the application of FTRANSLATE for generation purposes are the essentially new components in our generator. They are described in detail in [15].

4. Problem areas

In this section we will outline progress made on the selection of determiners (especially pragmatic anaphora) and the generation of paraphrases.

4.1 Selection of determiners

The suitability of determiners for appropriately expressing the role of an NP in a given constellation has not been treated very extensively. One of the few approaches in this direction is [9]. The aim there is mainly to produce unambiguous sentences in cases where scoping plays a crucial role. For instance, the choice between 'each' and 'all' ('a' and 'the same') is particularly stressed in the determiner selection to achieve scope expression reinforcement. We, however, focus on the choice of specificity and on the creation of fluent, possibly locally ambiguous sentences which can be interpreted in the context of a complete dialog.

In straightforward approaches, the number and the specifier features of NPs are direct derivations of the cardinality of the associated objects and of the fact that they have been mentioned earlier in the discourse. However, there are many instances which deviate from this standard pattern. As for the determination of the number feature, there is plenty of evidence that, apart from the cardinality of the referred object, *scoping* of the complete utterance plays a significant role. Hence, there are cases where a set of objects can be referred to by an NP in either singular or plural, depending on actual scoping.

In IRS-formula (1) E-EV is a quantifier for events (the buying event *z*) and the quantifier E-encodes the cardinality (of bond *y*) and the fact that *y* has not been mentioned earlier in the dialog. The term 'IBM' constitutes a simplification (it denotes the organization named 'IBM' and is treated as a constant here) and the time-intervals associated with states and events (which are the source for the determination of tense features) are omitted here. From formula (1) our verbalization component is able to generate, among some other possibilities, structures corresponding to the clauses (2) and (3), which are unconnected at this point.

- (1) ((3 x (MAN x))
 ((E- y (BOND y))
 (AND (HAS-ISSUER y 'IBM')
 ((E-EV z (BUYING z))
 (AND (HAS-AGENT z x)
 (HAS-THEME z y))))))
- (2) Three men bought a bond.
- (3) IBM has issued bonds.

When putting the clauses together, the first one is chosen to precede the second one on the surface level because 'man' has the widest

scope (this criterion may be overruled by focus constraints). As for the choice of determiners, only those for the 'bonds' are of real interest (in both clauses). In clause (2), the 'bonds' are within the scope of the 'men'. Hence, it is feasible to use either singular (corresponding to the quantifier) or plural (corresponding to the cardinality, which is derivable from the formula or, even simpler, can be obtained by a look-up in the A-Box). In actual conversation it seems more natural to use singular, perhaps because the singular form is less ambiguous than the plural. The plural variant is also vague with respect to the number of bonds the men bought (each or together). In our system, we select the singular variant unconditionally although a verification of the degree of precision might be achieved, for instance, by means of an anticipation feed-back loop which is used in HAM-ANS [17] for similar purposes.

As the resulting sentence, if considered without context, still has multiple readings, a disambiguating 'each' (or 'together' for reversed scoping) could be inserted. But, because we are dealing with a dialog system we prefer the more natural though locally ambiguous wording and trust the overall context without completely checking it. A comparable strategy is used by the analysis component: the ambiguity of an utterance is tolerated without asking for clarification as long as ordinary processing can continue despite the lack of precise information.

In clause (3), however, scoping of the 'bonds' is different (the 'men' are not present) which alters the choice entirely: the number feature must reflect the cardinality of the 'bonds' here. A more-or-less straightforward pronominalization and a passive transformation triggered by focus constraints lead to the sentences (4a) and (4b). Our speculation on contextual help has been immediately rewarded in this case: the second sentence provides the appropriate context for a unique interpretation of the first one.

- (4a) Three men bought a bond.
- (4b) They were issued by IBM.

By means of the components NORMALIZE and NORMALIZE-1 [16] HAM-ANS is able to handle sentence where scope reordering between the surface form and the underlying logic formula is involved. This includes also sentences like (4a) and (4b), but they can be treated only separately. In our approach the conceptual content can be expressed in a single formula and the verbalization procedure can select among possible surface expressions in a flexible way.

Formula (1) is also a good example for cases where dominance between NPs cannot be properly expressed in a single sentence ([9] gives criteria to detect such situations). Because of the different number features of the 'bonds' in sentences (4a) and (4b) our generator prefers to produce two separate sentences instead of embedding the second one starting with 'each of which ...'.

As for the choice of specificity, the simple approach mentioned earlier seems to work (partially) for objects in and of themselves. If, however, functional relations are involved, the uniqueness of the relation seems to play a similar role. This consideration refers to terminologically caused uniqueness as in phrase (5) (which is derivable from the associated number restriction defined in the T-Box - a bride can only have one father) and to uniqueness on the level of instances as in phrases (6) or (7).

- (5) the *father* of the bride
- (6) the *brother* of the bride
(if she has only one)
- (7) a *brother* of the bride
(if she has several brothers)

In cases like (7), knowledge of the speaker and not of the hearer is the decisive (and sufficient) factor. The hearer's knowledge can be augmented by the speaker's choice of determiner. A-Box knowledge is perfectly adequate here, but some care is necessary. Therefore, concepts and roles are annotated with meta-predicates INCONSISTENT, COMPLETE, INCOMPLETE, and UNKNOWABLE (as described in [1]) to avoid presupposition failures if the hearer's knowledge is more accurate than the speaker's.

Additional care is advisable if measurements are involved. According to regularities we have observed, an NP expressing a relation referring to a measurement requires the head NP to bear specifier feature INDEF even though the relation is unique whereas an NP expressing the same relation requires specifier feature DEF when joined with the object bearing it (compare phrases (8) and (9), to illustrate the difference).

- (8) The investment has a term of five years.
- (9) The term of the investment amounts to five years.

Thus, the T-box knowledge that an investment can have only one term is not sufficient. But, thanks to our detailed ontology, we can clearly recognize when a noun has been derived from a relation (all of which are represented by roles).

4.2 Pragmatic anaphora

In addition, our ontology helps us to determine more clearly the focus of attention, which is responsible for the validity of pragmatic anaphoric references. When an eventuality is mentioned in the dialog, all persons and objects involved (the fillers of the deep case roles) as well as their measurable properties are added to the focus of attention. Hence, sentence (11)

- (10) I want to invest my money.
- (11) What term should the investment have?

easily follows sentence (10) in a conversation, even though the 'investment' itself has not been previously mentioned. Moreover, the choice of the mood is remarkable in the previous sentence. In this case it is triggered by the (task-specific) assumption that the associated (consultation) object is a priori not identified. This assumption is maintained until an identifiable feature (e.g., the issuing number) is established for this object.

In the dialog control component [10] we have, among other things, incorporated inferences about simple state transitions to obtain, for instance, evidence about the effect of events. Thus, when the occurrence of an event is mentioned (e.g., as in sentence (12)),

- (12) Ich habe Geld geerbt.
I have inherited money.

the consequence of this event is also referable in the subsequent conversation. Thus, the possession of the money resulting from the inheritance might be referred to as '*der Besitz*' (the possession) in a subsequent system utterance. We are not sure how far we can go in this direction, but we believe that, to the extent that the inferential knowledge of the system is shared by the dialog partner, the creation of a pragmatic anapher is justified in such cases.

4.3 Generating paraphrases

One of the few approaches in this direction is the system EPICURE [8] which focuses on the generation of expressions which refer to objects whose quantities and shapes are crucial and may be subject to quick changes, integrating knowledge about the discourse structure as well (which we did not do in our system) to constrain the set of potential discourse referents. For instance, EPICURE is able to refer to a discourse entity at different stages of its existence in a single sentence to describe a shape change which the object referred to undergoes, like in '*Cut the onion into pieces.*'

Our emphasis lies on exploiting properties of objects as well as inferential knowledge to create expressions referring to objects in more indirect ways. An earlier approach, which is more comparable to ours, was taken in the VIELANG system [11]. Paraphrases are created primarily on the conceptual level, leaving the decision of whether or not to use one of them to subsequent processing. In that approach a paraphrase can refer to an entity by a superclass, by a role pertaining to the entity, or simply by a reduced form of a description previously used. Our method is more flexible distributing the burden of actually creating a paraphrase between FTRANSLATE, the NP-generation, and the verbalization component so that the decisions involved can be made at the most appropriate stage.

There is a rich variety for creating paraphrases in WISBER, where each of the subprocesses involved plays a particular role. This can be demonstrated by sentence (13) which results from a context substitution of the primary content specification (the state referred to by IRS-expression (14)) followed by substantial modifications in the course of the subsequent generation process.

- (13) Möchten Sie während der Laufzeit auf den Betrag zurückgreifen können?
Do you wish to have access to the sum [invested] during the term of the investment?
- (14) (LAMBDA (x)
 ((DS x (INVESTMENT x))
 (HAS-LIQUIDITY x HIGH)))

After substituting IRS-expression (14) into the user's WANT-context the result can be paraphrased by 'Should the liquidity of the investment be high?' The production of sentence (13) has been described in detail elsewhere [12]. In this context, we will concentrate on the motivation for the paraphrase 'der Betrag' (the amount) for the '' object. Originally, only the 'INVESTMENT' predicate is specified in the conceptual structure (14). This is also the case after the terminological transformation alters the role 'HAS-LIQUIDITY' into a complex expression signifying the 'possibility of having access to the money'. During NP-generation, the description of the 'investment' is expanded to include a disambiguating quantity of money because the investment otherwise might be confused with an object mentioned earlier in the dialog. In the subsequent verbalization process, the mappings of the 'investment' and the 'quantity of money' expressing its value collapse into a structure corresponding to the NP 'der Be-

trag', once again creating ambiguity. Again, a locally arising ambiguity is tolerated if it is resolvable in the dialog context.

Additionally the SUBSTITUTE schemata in our verbalization component provide us with means to immediately relate an NP containing a measurement (e.g., 40,000 DM) to two SEMs (the object quantified by the measurement and the measurement itself). Consequently, there is not a 1:n relation between SEMs and REFOs (as might be assumed intuitively), but rather an m:n relation. Hence, there is no problem in generating either of the sentences (16) or (17) as the successor utterance of sentence (15).

- (15) I have inherited 40,000DM.
 (16) That's a lot of money.
 (17) That's is a round number.

The same mechanism generally can be applied to a property whose description can be substituted for the object it belongs to like, for instance, the name of a person.

5. Conclusion

In this paper, we have briefly considered some aspects of a knowledge-based approach to generating referring expressions in a natural language advisory system. This approach combines conceptually (rather than lexically) based knowledge representation, semantic and pragmatic processes and syntactic information to provide a multipronged "human-like" attack on the problem of reference generation. Although constraints on time and resources limited the scope and coverage of our work, we were able to establish a base from which we hope to expand in future projects. Both our successes and the many unsolved problems we encountered in the course of our work lead us to the ineluctible conclusion that few of the problems of reference generation are likely to be solved unless all of the available resources of a dialog system are mobilized from the outset.

Acknowledgements

We would like to thank all our colleagues in the WISBER project for their contribution in the design and implementation of the system and for fruitful discussions in all phases of our work. In particular we indebted to Henning Bergmann, who designed and implemented FTRANSLATE, and to Heinz Marburger, who contributed the component dedicated to the generation of conceptual descriptions.

References

- [1] H. Bergmann, M. Gerlach: *Semantisch-pragmatische Verarbeitung von Äußerungen im natürlich-sprachlichen Beratungssystem WISBER*, in Wissensbasierte Systeme - GI-Kongreß 1987, W. Brauer, W. Wahlster (eds.), pp. 318-327, Springer (publ.), Berlin, 1987. Also in WISBER-Report Nr. 15, University of Hamburg, 1987.
- [2] H. Bergmann: *Short Description of FTRANSLATE*. WISBER Memo Nr. 30, University of Hamburg, 1987.
- [3] H. Bergmann, M. Fliegner, M. Gerlach, H. Marburger, M. Poesio: *IRS - The Internal Representation Language*. WISBER Report Nr. 14, University of Hamburg, 1987.
- [4] H. Bergmann, M. Gerlach: *QUIRK - Implementierung einer TBox zur Repräsentation begrifflichen Wissens*. WISBER Memo Nr. 11, second augmented edition, University of Hamburg, 1987.
- [5] R. Block: *Papers on Reference and Knowledge Representation*, WISBER Report Nr. 20, University of Hamburg, 1987.
- [6] R. Block: *Generating referential Expressions*. WISBER Report Nr. 46, University of Hamburg, 1989.
- [7] R. Brachman, J. Schmolze: *An Overview of the KL-ONE knowledge representation system*. Cognitive Science 9(2), pp. 171-216, 1985.
- [8] R. Dale: *Generating Referring Expressions in a Domain of Objects and Processes*. PhD Thesis, Centre for Cognitive Science, University of Edinburgh, 1989.
- [9] P.-J. Gailly: *Expressing quantifier scope in French generation*. In Proc. COLING-88, Budapest, 1988.
- [10] M. Gerlach, H. Horacek: *Dialog Control in a Natural Language System*. In Proc. EACL-89, Somers H., McGee M. (eds.), Manchester, 1989.
- [11] H. Horacek, E. Buchberger: *Achieving Text Coherence in a Generator for German Texts*. In Cybernetics and Systems '86, R. Trappl (ed.), pp. 831-836, Reidel (publ.), 1986.
- [12] H. Horacek et al.: *From Meaning to Meaning - A Walk Through WISBER's Semantic-Pragmatic Processing*. In GWAI-88, Geseke, W. Hoepfner (ed.), pp. 118-129, Springer (publ.), Berlin, 1988. Also in WISBER Report Nr. 30, University of Hamburg, 1988.
- [13] H. Horacek, C. Pyka: *Towards Bridging Two Levels of Representation Linking the Syntactic Functional and Object-Oriented Paradigms*. In International Computer Science Conference '88 - Artificial Intelligence: Theory and Applications, Hong Kong, J.-L. Lassez, F. Chin (eds.), pp. 281-288, December 1988. Also in WISBER Report Nr. 32, University of Hamburg, 1988.
- [14] H. Horacek: *Towards Principles of Ontology*. In GWAI-89, Geseke, D. Metzger (ed.), pp. 323-330, Springer (publ.), Berlin, 1989.
- [15] H. Horacek: *The Architecture of a Generation Component in a Natural Language Dialog System*. Appears in Current Research in Natural Language Generation, R. Dale, C. Mellish, M. Zock (eds.), Academic Press, 1990.
- [16] A. Jameson: *Documentation for Three HAM-ANS Components: Ellipsis, NORMALIZE and NORMALIZE-1*. HAM-ANS Memo Nr. 4, University of Hamburg, 1981.
- [17] A. Jameson, W. Wahlster: *User Modelling in Anaphora Generation: Ellipsis and Definite Descriptions*. In Proc. ECAI-82, pp. 222-227, 1982.
- [18] H.-J. Novak: *Generating Referring Phrases in a Dynamic Environment*. In Advances in Natural Language Generation, M. Zock, G. Sabah (eds.), Vol. 2, pp. 76-85, Pinter (publ.), 1988.
- [19] M. Poesio: *The QUARK Reference Manual*. WISBER Memo Nr. 22, University of Hamburg, 1988.
- [20] M. Sprenger: *Interpretation von Modalverben zur Konstruktion von Partnermodelleinträgen*. WISBER Memo Nr. 18, University of Hamburg, 1988.
- [21] I. Steinacker, E. Buchberger: *Relating Syntax and Semantics: The Syntactico-Semantic Lexicon of the System VIE-LANG*. In Proc. EACL-83, Pisa, Italy, 1983.