

On Inference-Based Procedures for Lexical Disambiguation

Jürgen Wedekind

Institute for Natural Language Processing

University of Stuttgart

Azenbergstr. 12

D-70174 Stuttgart, FRG

juergen@ims.uni-stuttgart.de

Abstract

In this paper we sketch a decidable inference-based procedure for lexical disambiguation which operates on semantic representations of discourse and conceptual knowledge. In contrast to other approaches which use a classical logic for the disambiguating inferences and run into decidability problems, we argue on the basis of empirical evidence that the underlying inference mechanism has to be essentially incomplete in order to be (cognitively) adequate. Since our conceptual knowledge can be represented in a rather restricted representation language, it is then possible to show that the restrictions satisfied by the conceptual knowledge and the inferences ensure in an empirically adequate way the decidability of the problem, although a fully expressive language is used to represent discourse.

1 Introduction

The determination of the contextual appropriateness of a reading of a lexically ambiguous sentence is commonly called *lexical disambiguation*. Lexical disambiguation presents a particular problem for any sort of natural language processing, but especially for machine translation, since the semantic grids of the source and the target may diverge in such a way that one has to disambiguate in cases where it is not required for applications of the source alone.¹ Resolving lexical ambiguities is a problematic task, since it involves different sources of linguistic and nonlinguistic information: information about the context of a sentence in a discourse, about the meanings of the words and about the world.

¹For the translation from German to English e.g. it is necessary to disambiguate 'Uhr' (clock/watch), but not for knowledge retrieval or other natural language processing tasks based on German alone.

Approaches to model the lexical disambiguation process formally differ as to the degree to which they consider the information of the various sources needed to disambiguate properly. We can distinguish two classes of approaches: "surface-oriented" approaches and "inference-based" approaches. Surface-oriented approaches rely on selectional restrictions (cf. e.g. McCord 1989) (sometimes supplied by an external type hierarchy/ontology (e.g. Nirenburg 1989) or are statistical (e.g. Kameyama, Peters, and Schütze 1993).

Although quite useful for some purposes, the performance of surface-oriented approaches is inherently limited in that their context sensitivity is always locally bounded (see e.g. Kay, Gawron, and Norvig 1994 for details). Since we cannot assume fixed finite context boundaries within each lexical ambiguity can be locally resolved, inference-based approaches seem more promising for handling lexical disambiguation. Inference-based approaches assume that the language of a logic is used to represent the meaning of a discourse, that the same language is used to store our conceptual and world knowledge and that resolution is achieved on the basis of the underlying logic by special inferences.

The most prominent inference pattern (which is also the center of the discussion here) is e.g. the proof of a contradiction from a given reading in a given context and our conceptual and world knowledge which allows us to rule out that reading. Although these approaches can handle the problem of disambiguating information arbitrarily far away (the whole context is available as a premise), without any further restrictions they run into tractability problems which exclude a practical application. Since we need -- as we will show below -- a representation language which is at least as expressive as the language of first-order predicate logic for an adequate representation of discourse meanings, an inconsistency test is not computable anymore if a classical (sound and as far as possible complete) calculus is used for the test: the underlying problem is simply undecidable.

Although it may turn out that the disambiguation problem is in fact undecidable if world knowl-

edge is also used for disambiguating inferences, we assume that resolution restricted to conceptual knowledge constitutes an important subproblem from a cognitive point of view, which is solvable in contrast to the general problem. That all known approaches which are confined to conceptual knowledge nevertheless run into problems is due to an empirically false estimation: they do not take into account that humans are able to disambiguate even without a full understanding of the discourse.² Lexical disambiguation works even very well in most of those cases where the discourse is inconsistent or its consistency is not known, and the inconsistency test would either fail or not necessarily terminate. Thus, the kind of reasoning which is involved in lexical disambiguation has to be essentially incomplete. Since our conceptual knowledge, on the other hand, can be represented in a rather restricted representation language, it is possible to restrict inferencing in an empirically adequate way which ensures decidability of the problem although a fully expressive language is used to represent discourse.

2 The Idea of Inference-Based Lexical Disambiguation

Lexical disambiguation is a procedure determining for a (lexically) ambiguous sentence within a discourse which reading of the sentence is contextually appropriate. From a logical point of view, the resolution of a lexical ambiguity is usually reconstructed by an inference process which rules out a reading if our conceptual knowledge contradicts this reading in the given context.³ In order to illustrate this type of inference-based resolution procedure let us consider the German sentence (1)

- (1) Einige Ärzte haben eine Schwester.

which contains the ambiguous lexical item 'Schwester'. Let us consider the two readings of (1) which have to be expressed in English by (2a,b).⁴

- (2) (a) Some physicians have a sister.
(b) Some physicians have a nurse.

These two readings are represented by the two (oversimplified) predicate-calculus formulas given in (3).⁵

- (3) (a) $\exists x(\text{Physician}(x) \wedge \exists y(\text{Sister}(y, x)))$
(b) $\exists x(\text{Physician}(x) \wedge \exists y(\text{Nurse}(y)))$

²Approaches which employ world knowledge in a nontrivial way are not known, by the way.

³An overview on the different methods is given in Kay, Gawron, and Norvig 1994.

⁴Sentence (1) has, of course, more readings. But we abstract away from the others for the sake of simplicity.

⁵Since we are primarily interested in the process, we abstract from further details, like temporal aspects.

Resolution of an ambiguity as in (1) is possible if it is embedded in a discourse which provides disambiguating information. If the discourse were continued as in (4)

- (4) Einige Ärzte haben eine Schwester, mit der sie verheiratet sind.

we could rule out the undesired reading given in (5).

- (5) $\exists x(\text{Phys}(x) \wedge \exists y(\text{Sister}(y, x) \wedge \text{Married}(x, y)))$

This reading which is expressed in English by (6)

- (6) Some physicians have a sister to whom they are married.

can be ruled out, since according to our conceptual system nobody can be married to his sister. Since this part of our conceptual knowledge can be formalized, as in (7)

- (7) $\forall x\forall y(\text{Sister}(y, x) \rightarrow \neg \text{Married}(x, y))$

the inappropriateness of reading (5) can be explicated from a logical point of view by the fact that we can derive a contradiction from that reading of (4) and our conceptual knowledge (meaning postulates).⁶

3 The Intractability Problem

Our inference-based reconstruction of the disambiguation process given in the previous section requires on the one hand that the meaning of the text is adequately represented in an appropriate (formal) representation language which allows the encoding of conceptual knowledge as well. By requiring on the other hand the underlying logic to be sound and as far as possible complete, we run, of course, into well-known decidability problems. Without any further restrictions on the expressive power of the representation language and/or the underlying logic the inconsistency of the representation of an arbitrary text and our conceptual knowledge is not decidable. Thus a natural language system whose resolver is based on such an inference system is not very useful, since an attempt to resolve an ambiguity is not guaranteed to terminate.

Since the field of AI which deals with knowledge representation and retrieval has been worrying about the same problem for quite a long time, it is not surprising that approaches to cope with

⁶There is, of course, another procedure which is dual to the given one. The dual variant allows us to rule out a reading if this reading of the discourse contains redundant information, i.e., information which already follows from the meaning postulates. This procedure would exclude e.g. for 'Einige Ärzte haben eine Schwester, mit der sie nicht verheiratet sind' the *Sister* reading which is expressed in English by 'Some physicians have a sister to whom they are not married', since (7) implies for physicians who have a sister that they are not married to her.

this problem within lexical disambiguation were directly adopted from knowledge representation. According to the subject of the restriction used to ensure the tractability of the problem, we have to distinguish three main approaches.

The simplest way to guarantee tractability of the disambiguation problem is by *restricted computations*. If the underlying logic of a resolver is known to be undecidable (e.g. the inference machine used in LILOG (Bollinger, Lorenz, and Pletat 1991)) the only chance to ensure termination is by stopping the computation after a limited amount of resources (inference length, computation time, etc.) is consumed. Since the termination behavior of such a system is without any further empirical evidence not in any way correlated with our cognitive capabilities and without any further formal evidence not in any way correlated with the behavior which we would expect if the disambiguation problem were nevertheless decidable, we have to rule out these approaches from a scientific point of view.

The second class of approaches achieves tractability by *restricted representation languages*. These restrictions allow one to base retrieval on a tractable logic which is sound and complete. In order to support the distinction between terminological and assertional knowledge, most formalisms of this class provide two different (restricted) representation languages: the terminological language and the assertional language.

To use one of these knowledge representation formalisms (especially the tractable descendants of KL-ONE) for lexical disambiguation leads to problems which disqualify language restrictions as the only means to ensure tractability of the disambiguation problem. On the one hand it is, of course, possible to find examples of meaning postulates which are inexpressible in the restricted terminological languages (see e.g. the list given in Doyle and Patil 1991). But these counterexamples do not provide conclusive arguments, since the expressive power needed in order to formulate these counterexamples is still rather weak, and one could counter by moving a little bit of expressive power around. Much more crucial for disambiguation are the restrictions imposed on the assertional language.

In BACK (Hoppe et al. 1993), for example, which is used by Quantz and Schmitz 1993 for disambiguation by storing the text representation in the ABox (assertional knowledge base) and the meaning postulates in the TBox (terminological knowledge base) it is e.g. not possible to represent (4) in an adequate way. We can only find representations whose models include the models of (5), but not a representation with exactly the same models. In order to see this, consider the set-theoretic versions of the satisfiability conditions of (5) and (7) (for a model with interpretation

function \mathcal{I}) given in (8) and (9).⁷

$$(8) ([Phys.]^{\mathcal{I}} \cap \{x \mid \exists y ((y, x) \in [Sist.]^{\mathcal{I}} \cap [Marr.]^{\mathcal{I}})\}) \neq \emptyset$$

$$(9) [Sister]^{\mathcal{I}} \cap [Married]^{\mathcal{I}} = \emptyset$$

According to these conditions the BACK expressions (10) and (11) were adequate representations of (5) and (7).

$$(10) X :: Phys. \text{ and some}(Sister \text{ and Married})$$

$$(11) Sister \text{ and Married} :< \text{nothing}$$

Although (10) contradicts the TBox representation (11) of (7), it is not possible to use BACK to establish this inconsistency (incoherence), since BACK does not allow the conjunction of roles in the ABox (cf. Hoppe et al. 1993, p. 50) which is of course needed in (10) (the conjunction of the roles *Sister* and *Married*).

Example (10) is, of course, just beyond the border of the permitted expressions, since it is in principle expressible but not allowed, and much more problematic (e.g. for ‘donkey’ sentences) is certainly the fact that variables are not explicitly available in these representation languages. But it should indicate the lack in expressive power at least inasmuch as it is possible without a more general formal proof (which we cannot give here for lack of space). Since the correct disambiguating inferences cannot be performed anymore if the truth conditions of a discourse are boiled down in a way that allows to represent it (somehow) in such a restricted assertional language, approaches which model lexical disambiguation on the basis of these knowledge representation formalisms must fail.

Since an extension of the expressive power of the assertional languages would lead immediately to our original tractability problem, we have to give up the implicit assumption that lexical disambiguation presupposes the consistency of the discourse, if we don’t want to give up lexical disambiguation at all. Thus, we end up in the third class of approaches which provide us with fully expressive languages to represent discourse and ensure tractability by *limited inferences*. In order to see whether the requirements of soundness and completeness can be adequately weakened we have to study the inferences involved in lexical disambiguation more carefully.

4 Towards Tractable Lexical Disambiguation

To limit inference is a well-known strategy employed for knowledge retrieval (e.g. Frisch and Allen 1982). By using incomplete theorem provers it is certainly possible to ensure tractability, but incompleteness is always a compromise which can

⁷We assume *Married* to be a symmetric relation.

be accepted as long as the prover computes the desired inferences completely (which is in fact hard to show).

In contrast to knowledge retrieval where incompleteness is assumed for utility reasons, inference systems used for lexical disambiguation have to be essentially incomplete. Otherwise we would get wrong results. In order to motivate our restrictions we proceed in three steps. In the first step we show that we need an incomplete (but sound) inference mechanism for lexical disambiguation, since a complete mechanism leads to wrong results. We specify a class of inconsistency proofs which contains the disambiguating inferences as a subclass. In the second step, we separate out those proofs which are in fact disambiguating and illustrate in the last step that the discourse structure imposes further restrictions on the accessibility of premises.

4.1 The Incompleteness and Decidability of Lexically Disambiguating Inference Mechanisms

In order to develop our approach to lexical disambiguation, we work successively through some adequacy conditions which have to be satisfied by an adequate procedure. According to the discussion in section 3 we have to assume a fully expressive language for the representation of discourse. Assumption (I) is therefore as follows:

- (I) We have to assume a fully expressive language for the representation of texts. Semantic representations of natural language texts in this language do in general not satisfy conditions which make them decidable (see e.g. Rabin 1977 for standard conditions).

To illustrate which kind of incompleteness we need, we assume that the meaning postulates and the discourse can be expressed in a first-order language without function symbols and identity. Although we think that one needs a more expressive language for an adequate representation of discourse, and that very often nonmonotonic reasoning is involved, the first-order case seems nevertheless representative, since we have to deal with the decidability problem. Moreover, we expect that the methodology we used can be applied to more expressive discourse representation languages in a similar way.

For our conceptual knowledge on the other hand we make the much stronger assumption (II).

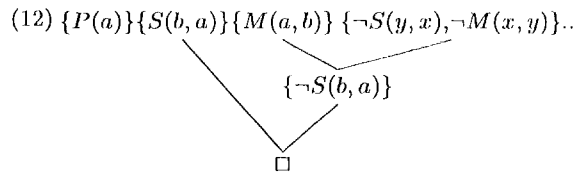
- (II) Conceptual knowledge is represented by a finite consistent and decidable set of meaning postulates MP that does not contain logically valid subsets of formulas.⁸

⁸Since this condition is certainly not satisfied by our world knowledge, its integration in the disambiguation process would be a much harder problem.

Decidability of MP, i.e. the decidability of $MP \vdash \phi$ for a given formula ϕ , results from the fact that MP does not make any absolute existential claim on the entities in the world, especially on their cardinality.⁹

In order to be able to specify the incompleteness of our inference machinery in terms of a resolution logic, let us in the following assume that MP and the discourse is given in *Skolem conjunctive form* (SCF). I.e., as two universally quantified formulas whose matrices are in conjunctive normal form. Let us furthermore assume that we would know that the given discourse is consistent (we abstract here first from the problem that this property is undecidable). We were then able to determine the unsatisfiability of the discourse and MP by resolution.

Let us take, for example, the set of clauses obtained from the SCF's of the meaning postulate (7) and the discourse (5) by the standard preparation procedures. If we abbreviate *Physician* by *P*, *Sister* by *S* and *Married* by *M* and use clause set notation (each conjunct of the matrix is represented as the set of its disjunctively connected literals) the unsatisfiability of (5) and (7) can be shown, since there is a resolution refutation depicted as a refutation tree in (12).



The whole problem is now that despite of the decidability of MP the lexical disambiguation problem would still be undecidable if it would presuppose a consistent discourse. Decidability of the lexical disambiguation problem results nevertheless from the fact that lexical disambiguation does not involve a complete understanding of the discourse. In order to illustrate that, let us

⁹By checking several examples we found out that this property can be characterized model-theoretically as follows. There is a finite set of (up to isomorphism unique) finite models $\{M_1, \dots, M_n\}$ of MP such that each other finite model M'_k of MP can successively be reduced to a model $M \in [M_k]$ by a chain of models $M = M_k^0 \prec M_k^1 \prec \dots \prec M_k^n$ of MP such that for each pair of models $M'_k = \langle \mathcal{U}^i, \mathfrak{S}^i \rangle$, $M_k^{i+1} = \langle \mathcal{U}^{i+1}, \mathfrak{S}^{i+1} \rangle$ there is a (partial) isomorphism f from $\mathcal{U}^{i+1} \setminus \mathcal{U}^i$ in \mathcal{U}^i such that $\mathfrak{S}^i(R)$ is the set of tuples $\langle a_1, \dots, a_m \rangle$ with $\langle b_1, \dots, b_m \rangle \in \mathfrak{S}^{i+1}(R)$, and $a_l = b_l$ if $b_l \in \mathcal{U}^i$, and $a_l = f(b_l)$ if $b_l \in \mathcal{U}^{i+1} \setminus \mathcal{U}^i$, for every relation symbol R . Since the infinite models of MP correspond to unions of infinite chains of such models (i.e. MP is a rather restricted $\forall\exists$ -theory), we can reduce the test of $M'_k \models \phi$ for each model M'_k of MP to a test of $M_k \models \phi$. Thus, we can decide $MP \vdash \phi$ by checking $M \models \phi$ for all models $M \in \{M_1, \dots, M_n\}$. But note that this does not allow us to test whether ϕ is valid or not.

consider the inconsistent lexically ambiguous sentences (13a,c) whose *Sister* readings are expressed in English by (13b,d).

- (13) (a) Es gibt keine Schwestern, aber einige Ärzte haben eine, mit der sie nicht verheiratet sind.
 (b) There are no sisters at all, but some physicians have one to whom they are not married.
 (c) Es gibt keine Schwestern, aber einige Ärzte haben eine, mit der sie verheiratet sind.
 (d) There are no sisters at all, but some physicians have one to whom they are married.

Although it is possible to derive from the semantic representations of (13a,c) a contradiction, these proofs are by no means disambiguating inferences, since the meaning postulates are not involved. In order to be able to explain by inconsistency proofs why the *Sister* reading is excluded for (13c) but not for (13a) one has to assume an incomplete inference system.¹⁰ Otherwise the system would not work correctly and would, of course, not necessarily terminate. Thus, our third assumption is:

- (III) Lexical disambiguation is very often possible although the discourse is inconsistent or its consistency is not known.

What we are in fact looking for is a procedure which tests whether there is a consistent set of information pieces of the discourse which contradicts MP. In order to isolate the consistent information pieces provided by a (possibly inconsistent) discourse we use a discourse representation (and meaning postulates) in clause form. Since each single clause of such a representation must be satisfiable, we can identify the set of consistent information pieces provided by a discourse with the set of clauses of the discourse in SCF. On the basis of this set we can then test whether there is a consistent subset of these pieces which contradicts MP. Take as an example the clause representation of (13a) and our meaning postulate (7) depicted in (14a,b).

- (14) (a) $\{\neg S(u, v)\} \{P(a)\} \{S(b, a)\} \{\neg M(a, b)\}$
 (b) $\{\neg S(y, x), \neg M(x, y)\} ..$

That the *Sister* reading is not excluded for (13a) is then explicable by the fact that there is no consistent subset of clauses of (14a) which is inconsistent with MP. What is consistently said in the (inconsistent) discourse does not violate the

¹⁰For the sake of simplicity we were confined to short and simple examples and could therefore not avoid some artificiality. Moreover, an additional test based on the procedure sketched in footnote 6 would certainly exclude the *Sister* reading for (13a). But it is, of course, easy to construct more realistic examples where the inconsistency is much more hidden and does not affect the disambiguation.

meaning postulates in this case. In order to test this kind of incompatibility we have to demand that each resolution deduction starts with a clause from MP. This restriction prevents the attempt to prove the inconsistency of the discourse alone (at least if MP does not contain logically valid subsets of formulas that we assume and are able to decide). It prevents us from proving the unsatisfiability of (14a,b), but we can still show the inconsistency of the clause representation of (13c) and (14b) as in (12).

4.2 Disambiguating Inferences

The restriction introduced above is by no means sufficient, since the proof procedure is not yet sensitive to the predicates representing the readings of an ambiguous lexical item. In order to illustrate this insufficiency let us consider the English translation of the *Sister* reading of (4), repeated in (15).

- (15) Some physicians have a sister to whom they are married.

If we also assume (7) for English then a contradiction would result although we did not regard 'sister' as ambiguous (at least in our oversimplified language domain). Hence, if (15) were embedded in a larger discourse we would have no chance to disambiguate other ambiguous lexical items, since we would get a contradiction for every reading of these items. That disambiguation is nevertheless possible in many of those cases can be made obvious e.g. by continuing (15) as in (16).

- (16) Some physicians have a sister to whom they are married. Some of these sisters admire stars who got an Oscar.

The disambiguation of the ambiguous item 'star' should make no problems, given we had the right meaning postulates. Thus, we have to assume:

- (IV) Lexical disambiguation is very often possible although the discourse contradicts our conceptual knowledge.

In order to disambiguate properly we have to consider only those consistent sets of information pieces which contain at least one occurrence of the predicate that represents one reading of the ambiguous lexical item. Therefore we have to demand in addition that each resolution deduction starts with a pair of clauses $A \in MP$ and B from the discourse representation where B contains an occurrence of the predicate representing one reading of the ambiguous lexical item. This prevents disambiguating inferences for cases where there is no choice with respect to the interpretation of the discourse ('sister' has to be interpreted as *Sister* although there is a contradiction).

4.3 Reflecting Discourse Structure

For lexical disambiguation we assumed so far that the underlying inference machinery operates on

the set of consistent information pieces provided by the discourse. This set was crucially dependent on what is said and not on what follows, since we were (especially in case of inconsistencies) not interested in the set of all logical consequences of a discourse. Hence, our procedure already reflects in a very weak sense the discourse structure, since we did not allow all conversions preserving logical equivalence, but only those needed to construct an SCF from the discourse.

By converting the whole discourse into SCF we made all consistent information pieces provided by the discourse accessible for lexical disambiguation. Whether we need this entire set or just a rather limited subset of pieces which can be made accessible by locally restricted conversions into SCF, is for a first-order discourse an empirical but no formal problem. But if we consider discourse representations in more expressive languages (e.g. the language of an intensional logic) it becomes clear that we have to make only those consistent pieces accessible which result from first-order consequences of the discourse representation. Information in the scope of the intensional verb in (17a) whose *Sister* reading is expressed in English by (17b) is, for example, not accessible for lexical disambiguation.¹¹

- (17) (a) Einige Ärzte versuchten ihre Schwestern zu heiraten.
 (b) Some physicians tried to marry their sisters.

Since we cannot get an SCF of the first-order consequences of a (possibly inconsistent) discourse represented in a more expressive representation language, it is necessary to find exactly those logical equivalence preserving conversions which allow us to convert the discourse representation in such a way that the adequate set of consistent information pieces can be made accessible for the disambiguation by locally restricted conversions into SCF. But we must, of course, admit that further study is needed in order to be able to determine these conversions.

5 Conclusion

Lexical disambiguation is a procedure which works according to the communicative convention to interpret the discourse as consistent as possible, if there is a choice. It allows us to decide for two alternative readings of the discourse which one is less contradictory to what is said consistently in the discourse and to our conceptual knowledge. As the analysis of the examples in this paper has shown, there is a striking similarity between lexical disambiguation and anaphoric resolution. Not

a complete understanding of the discourse is required, but only an incomplete one that is restricted to a set of accessible consistent information pieces. The only difference is that lexical disambiguation requires a little bit more understanding.

Acknowledgments

Thanks to Ede Zimmermann and Hans Kamp for useful discussions and to the anonymous reviewers for comments.

References

- Bollinger, T., S. Lorenz, and U. Pletat. 1991. The LLOG Inference Engine. In O. Herzog and C. Rollinger, eds., *Text Understanding in LLOG*. Berlin: Springer.
- Doyle, J., and R. Patil. 1991. Two Theses of Knowledge Representation: Language Restrictions, Taxonomic Classification, and the Utility of Representation Services. *Artificial Intelligence* 48, 261-297.
- Frisch, A., and J. Allen. 1982. Knowledge Retrieval as Limited Inference. In D. Loveland, ed., *6th Conference on Automated Deduction*. Berlin: Springer.
- Hoppe, T., C. Kindermann, J. Quantz, A. Schmiedel, and M. Fischer. 1993. BACK V5 Tutorial & Manual. KIT Report 100. TU Berlin.
- Kameyama, M., S. Peters, and H. Schütze. 1993. Combining Logic-Based and Corpus-Based Methods for Resolving Translation Mismatches. In *Working Notes: AAAI Spring Symposium Series, Symposium "Building Lexicons for Machine Translation"*, Stanford.
- Kamp, H. 1992. Disambiguation in Discourse. In M. Aurnague et al., eds., *Semantics of Time, Space and Movement*. Working Papers on the 4th International Workshop, Chateau de Bonas, France.
- Kamp, H., and A. Roßdeutscher. 1994a. Remarks on Lexical Structure and DRS-Construction. *Theoretical Linguistics* 20, 98-164.
- Kamp, H., and A. Roßdeutscher. 1994b. DRS-Construction and Lexically Driven Inference. *Theoretical Linguistics* 20, 166-235.
- Kay, M., J. Gawron, and P. Norvig. 1994. *VERBMOBIL: A Translation System for Face-to-Face Dialog*. Stanford: CSLI Lecture Notes.
- McCord, M. 1989. Design of LMT: A Prolog-Based Machine Translation System. *Computational Linguistics* 15, 33-52.
- Nirenburg, S. 1989. Knowledge-Based Machine Translation. *Machine Translation* 4, 5-24.
- Quantz, J., and B. Schmitz. 1993. Knowledge-Based Disambiguation for Machine Translation. *Minds and Machines*. Special Issue on 'Knowledge Representation for Natural Language Processing'.
- Rabin, M. 1977. Decidable Theories. In J. Barwise, ed., *Handbook of Mathematical Logic*. Amsterdam: North Holland.

¹¹More complex examples can be found e.g. in Kamp 1992, Kamp and Roßdeutscher 1994a,b.