NATURAL LANGUAGE INFORMATION RETRIEVAL SYSTEM DIALOG

L. Bolc, K. Kochut, A. Leśniewski[1], T. Strzałkowski[2]

Warsaw University
Institute of Informatics
PKiN, pok.850
00-901 Warszawa
P O L A N D

ABSTRACT

Presented paper contains a description of an experimental version of the natural language information retrieval system DIALOG. The system is destined for the use in the field of medicine. Its main purpose is to ensure access to information to phisicians in a conversational manner. The use of the system does not require ability of programming from its user.

1. Introduction

The paper presents the state of elaboration of the natural language information retrieval system DIALOG. Its aim is an automatic, conversational extraction of facts from a given text. Actually it is real medical text on gastroenterology, which was prepared by a team of specialists. The system has a modular structure.

The first, and in fact very important module is the language analysis module. Its task is to ensure the transition of a medical text from its natural form, i.e. sentences formed by physicians, into a formal logical notation. This logical notation, i.e. logical formulae, is rather universal and can be easy adapted to various deductive and knowledge representation methods. The program of the analyser was written with the use of the CATN /Cascaded ATN/ technique, where the syntactic and semantic components constitute separate cascades.

In the deduction and knowledge representation module the weak second order language was used. The works by E.Konrad /Konrad 76/ and N.Klein /Klein 78/ from the Technical University in Berlin were the starting point in the elaboration of this module.

Presented version of the system was implemented on the IBM 370 computer /VM 370 operating system/.

2. Transformation of natural language sentences into logical formulae

The user of the DIALOG system introducing his utterance into the system comes into direct contact with the natural language analysis module. This module plays the key role in the machine natural language communication process. Similarly as in many other information systems of this type, e.g. LUNAR /Woods 72/, PLANES /Waltz 76/, SOPHIE /Burton 76/, RENDEZ-VOUS /Codd 78/, PLIDIS /Berry-Rogghe 78/, DIALOGIC /Grosz et al. 82/, the purpose of the module is to transform a text in the natural language into a chosen formal representation. Such Such a representationmust meet a number of requirements. Firstly, it must be "intelligible" to the internal parts of the system, i.e. the deductive component and/or managing the data base. Secondly, it must carry in a formal and clear maner the sense and meaning of utterances in natural language. Finally, the representation should allow for a reproduction of the original input sentence with the aim of generating intermediate paraphrases and/or answers for the user.

In the parser of the DIALOG system, we attempted on the gratest, in our opinion, achievements in the field of natural language processing. The following works had the greatest influence on the final form of the module: /Berry--Rogghe 78/, /Bates 78/, /Carbonell 81/, /Cercone 80/, /Chomsky 65/, /Ferrari 80/, /Fillmore 68/, /Gershman 79/, /Grosz 82/, /Landsbergen 81/, /Marcus 80/, /Martin 81/, /Moore 81/, /Robinson 82/, /Rosenschein 82/, /Schank 78/, /Steinacker 82/, /Waltz 78/, /Wilensky 80/, /Woods 72/ and /Woods 80/. We have transferred, with greater or less success, the most valuable achievements presented in these works, pertaining

---

1 Presently:
Universitat Stuttgart, Institute fur Informatic, Herdeweg 51 Raum 4, Postfach 560.7000 Stuttgart, FRG

2 Presently:
Simon Fraser University, Dept. of Computing Sci., Burnaby, B.C. Canada

mainly to the English language processing, into our system, using them in the treatment of the Polish language. We attempted thus, to preserve a certain distance with regard to the language itself, as well as the subject of conversation with the computer, so that the adapted solutions were of a broader character and through that became comparable with the state of research in that field in other countries.

2.1. The role, place and structure of the language analysis module

The purpose of the language analysis module in the DIALOG system is transformation of the user´s utterance /in Polish/ into the I order logic formulae. Other formal notations such as II order logic formulae, FUZZY formulae, Minsky frames and even the introduction of intensional logic elements are also considered. At present, we will concentrate on the process of transforming a natural sentence into a I order logic formula.

The system is equipped with two independent modules: deduction and data base management. The data for these modules are the formulae generated by the parser. We will present only one module working on the basis of the weak second order logic.

The parsing system consists of the two closely cooperating parts: a syntactic analyser and a semantic interpreter. The whole was programmed with the aid of a mechanism called CATN /Cascaded ATN/ /Woods 80/, /Bolc, Strzalkowski 82a,82b/ /Kochut 83/, where the syntactic component plays the role of the "upper", i.e. the dominating "cascade". For the syntactic analyser produces a structure of the sentence grammatical analysis, which in turn undergoes a semantical verification. In case, where the semantic interpreter is not able to give the meaning of the sentence, the syntactic component is activated again with the aim of presenting another grammatical analysis. If such an analysis cannot be found, the input sentence is treated as incorrect.

2.2. The syntactic analyser

The syntactic component of the parser produces a grammatical analysis of the input sentence in Polish. This was possible due to a skillful programming of rules governing the morphology and syntax of the language. Although, the whole system was oriented towards a defined type of texts /medical/, the accepted solutions make it a much more universal tool. We do not claim that the syntactic analyser in its present form is able to solve all or the majority of problems of the Polish language syntax. It includes, however, rather wide subset of the colloquial language, enriched by constructions characteristic for medical texts.

A natural language sentence introduced into the parser undergoes firstly a pretreatment in a so called spelling corrector. If all the words used in the sentence are listed in the system vocabulary then the sentence is passed for syntactic analysis. Otherwise the system attempts to state whether the speaker made a spelling error, giving him a chance to correct the error and even suggesting the proper word, or whether he used a word unknown to the system. In the last case, the user has a possibility of introducing the questioned word into the vocabulary but in practice it may turn out to be too troublesome for him. Usually then, the user is given a chance of withdrawing the unfortunate utterance or formulating it in a different way.

The proper syntactic analysis begins at the moment of activating the first "cascade" of the parser. It consists of five ATN nets, with the aid of which the grammar of the subset of the Polish language has been written. The two largest nets SENTENCE /sentences/ and NOUN-PHR /nominal groups/ play a superiorrole in relation to others: ADH-PHRA /adjective groups/, ADV-PHRA /adverb groups/ and Q-EXPR /question phrases/. The process of syntactic analysis is usually quite complex and uses essentially the non-deterministic character of processing in ATN. It is justified by the specific nature of the Polish language, which is characterised by a developed inflection and a sentence free word order.

The result of the syntactic analysis is a grammatical analysis of the input sentence in the form of a so called o-form. It is a nonflexional form of a sentence, ordered according to a fixed key. The construction of the o-form can be expressed ba the structure:

```
<o-form>  ::=
 (S <questions> | <negation> | <modalitie>
|<predicate/verb> | <vague> | <subject> |
 <direct object> | <indirect object> |
{<prep. phrase>|}* {CAUSE/RESULT <o-form>}*
  END)
```

The stick mark "|" is usually used as a symbol of the meta-language. Here it is used as a symbol of the defined language. Symbols S and END comprise a single clause. A clause expresses every elementary activity or event expressed in the input sentence. Often, the o-form has a richer structure than a classical analysis tree. The elements of the o-form called <subject> , <direct object>, <indirect object>, and <adjective phrase> can also be expressed or modified with the use of clauses. The stick marks "|" separate the parts of the o-form and are its constatnt elements. Then transformed question is subjected to semantic interpretation.

The syntactic analyser manages the vocabulary, where inflexional forms of words are kept. The vocabulary definition specifies the syntactic categories, to which given words belong. It also describes forms of words with the aid of lexicalparameters: case, number, person and gender. These parameters are of gret value in examining the grammatical construction of sentences.

2.3. The semantic interpreter

When the syntactic analysis is successfully completed the o-form of the input dentence is forwarded for the semantic interpretation. The syntactic "cascade" is suspended, i.e. removed from the operational field, leaving place for the semantic "cascade". The configuration of the removed "cascade" is remembered thus, in case of necessity of generating an alternative grammatical analysis.

The semantic interpreter consists of the two main parts: a constant controlling part, working on the basis of a very general pattern adjustment, and compatible experts algorithms, where the knowledge of the system in the field of conversation has been coded. The process of interpretation is assisted by a special vocabulary of semantic rules and on additional vocabulary complementing the expert knowledge.

The sentence in the o-form is forwarded directly to the controlling part of the interpreter, where such its parameters as time, negation, aspect .... are evaluated first. Then the central predicative element of the sentence "calls for" a proper semantic rule, which from then will guide the interpretation process. The rule has a form of a pattern-concept pair /Wilensky 80/ , /Gershman 79/, /Carbonell 81/, where the pattern reflects the scheme of an elementary event, wheras the concept indicates how its meaning should be expressed through formulae. The semantic rule is activated for the time of interpretation of a single clause. If the pattern is adjusted to the clause, an atomic formula is generated, expressing the meaning of the clause. The meaning of the whole sentence is expressed as a logical combination of meanings of all the o-form clauses. The semantic rules bring different /on the surface/ descriptions of the same phenomenon into a common interpretation.

The general structure of formulae generated by the interpreter is expressed by an implication:

$$\alpha_1 \wedge \alpha_2 \wedge \ \ldots \wedge \alpha_n \longrightarrow \beta$$

where β has been introduced from a semantic rule and α_i come from the system knowledge - special compatible parts of the interpreter called the experts. Individual o-form phrases, in the context of the dialogue subject, are interpreted in experts.

In our system, designed for conversation with a phisician, we have experts for names of sicknesses /SICKNESS/, names of organs /ORGAN/, internal substances /SUBSTANCE/, therapies /TREATMENT/, medicaments /MEDICAMENT/ and names of animate objects /ANIMATE/ and the remaining objects foreign to the body /PHYSOBJ/. Experts are activated on the request of a proper semantic rule. The controlling part of the interpeter "instructs" the expert/s/ chosen by the pattern to interpret a notion or expression. The indicated expert can solve the problem on its own or seek for the help of other experts. Often, one complex expression has to be qualified by two or three exprrts.

All the experts, as well as the controlling part of the interpreter /FORMULA, CASES and QWORDS nets/ have been recoreded in ATN formalism and form a lower "cascade" of the parser.

The interpreter is also equipped with a mechanism of context pronominal reference solution.

2.4. Examples of transformation of a medical text into logical formulae

We will present two examples of transformation of medical sentences into I order logic formulae. Before that, a few words on the adopted convention of formula notation. The symbols IMPLSYM and KONJSYM are logical operators $\longrightarrow$ /implication/ and $\forall$ /conjunction/ respectively. Integer placed directly after the symbol KONJSYM indicates the number of conjunction factors. Names of predicates are preceded by symbols "#" /hash mark/, and an integer placed right to the name defines the number of predicate arguments. The arguments specify their type /sort/, name of the variable and constant /if there is one/.

Example 1

Sentence:

Alkohol powoduje również wzrost napięcia mięśniówki dwunastnicy.

/Alcohol also causes the rise of the tonicity of the duodenum muscular coat/

o-form:

```
(S DCL |||| POWODOWAC | ROWNIEZ |
ALKOHOL | S |||| WZROST ||| NAPIECIE
MODIFIERS MIESNIOWKA DWUNASTNICA
||| END ||| END)
```

formula:

```
(IMPLSYM
  (KONJSYM 3 ((#BADMEDIC 1) (MEDIC X44))
    ((#MEDICAMENT 2) (MEDIC X44) (MNAME
                X45 ALKOHOL))
    (IMLSYM
      (KONJSYM 4((#WYDZ-NARZAD 1)
                    (ORGAN X49))
        ((#ORGAN 2) (ORGAN X49) (ONAME
                    X50 DWUNASTNICA))
        ((#PART-OF-ORGAN 3) (BODY X48)
              (PNAME X51 MIESNIOWKA)
              (ORGAN X49))
        ((#SICKNESS 4) (SICK X47)
              (STYPE X52 FIZJ)
              (SNAME X53 NAPIECIE)
              (BODY X48)))
      ((#RISE 2) (SYMPTOM X46) (SYMPTOM X47)
 ((#IMPLY 3) (INFER X43) (MEDIC X44)      )))
          (SICKNESS X46)))
```

Example 2

Sentence:

Czy alkohol może być przyczyną OZT?

/Can alcohol be the cause of acute pancreatitis ?/

o-form:

```
(S CZY || MOC || BYC || ALKOHOL |
PRZYCZYNA MODIFIERS OSTR* ZAPALENIE
TRZUSTKA ||| END)
```

formula:

```
(NIL (IMPLSYM
    (KONJSYM 6((#VAGUE 2) (ACTION X69)
                  (VAG X70 MOC))
      ((#BADMEDIC 1) (MEDIC X71))
      ((#MEDICAMENT 2) (MEDIC X71)
            (MNAME X72 ALKOHOL))
      ((#ORGAN 2) (ORGAN X74)
            (ONAME X75 TRZUSTKA))
      ((#WYDZ-NARZAD 1) (ORGAN X74)
      ((#SICKNESS 4) (SICK X73) (STYPE
            X76 PATO) (SNAME X77 OZT)
            (BODY X74)))
    ((#IMPLY 3) (INFER X69) (ETIO X71)
            (SICKNESS X73))))
```

3. The deduction and knowledge representation module

The deduction module is a separate part of the whole DIALOG system. Its main purpose is to collect and represent the knowledge gained by the system and also the ability to use the possessed information in accordance with the wishes of the user of the system.

Our work on the achievement of the objectives indicated above was based on the experiences presented by E.Konrad and N.Klein /Konrad 76/, /Klein 78/ from Technical University in West Berlin.

In the previous chapter we presented how the text, written in Polish, is transformed into I order logic formulae. This, of course, implies the way of representation of the knowledge presented in the natural language.

3.1. Knowledge representation

The information included in the logical formulae coming from the language module has to be stored for later use. The logical formulae are then introduced into the data base. The data base, adequately filled with the mentioned formulae, constitutes the knowledge representation carried through the natural language sentences. It is as equivalent to the text as the I order logic allows to convey the meaning of the natural language sentences.

Data Base

The date base consists of three separate parts: a nucleus, an amplifier and a filter /Konrad 76/. Each of the parts includes a different , from the conceptional point of view, elements:

A. The nucleus includes groud literals, which represent facts occuring in the field of knowledge represented in the base. E.g.the information that the pancreas is a secretory organ is presented as a literal

(# WYDZ-NARZAD (TRZUSTKA))

From the system point of view there is no conceptional difference between the teo facts: the above one,and

(ORGAN (TRZUSTKA))

Thus the type /sort/ ORGAN may be regarded as a predicate and the above atomic formula as true one.

B. The amplifier is a part representing the "fundamental" knowledge of the system. The formulae included in the amplifier can be devided into three categories:

1/ dependent formulae

/i/ $\forall x_1[s]..\forall x_n[s]\ A[x_1,\ldots,x_n]\to\Pi(x_1,\ldots,x_n)$

A is here any formula and Π a predicate. As we can see each variable, bound by the universal quantifier is of a specified sort.

2/ independent formulae

/ii/ $\forall x_1[s]\ldots\forall x_n[s]\ \Pi(x_1,\ldots x_n)$

3/ restrictive formulae

/iii/ $\forall x_1[s]\ldots\forall x_n[s]\ \neg\Pi(x_1,\ldots,x_n)$

The majority of the formulae generated by the language analysis module is of the /i/ form.

C. The filter contains the formulae representing the knowledge necessary to preserve the integrity of the data base.

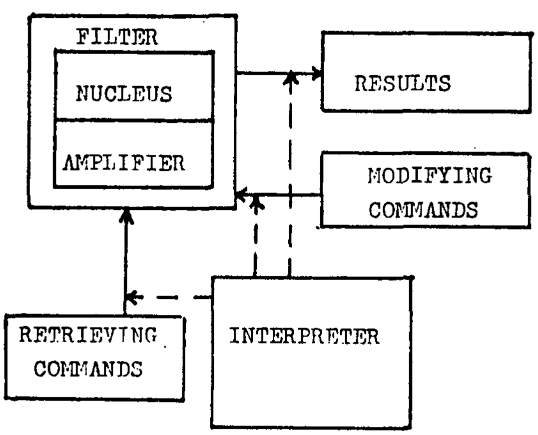

Fig. 1. Diagram of the data base system /Konrad 76/

Recapitulating, the nucleus represents the extensional part of the knowledge represented in the data base. It is the fundamental knowledge which cannot be obtained from the analysis of the presented text, and which is assential to proper deduction. The amplifier represents the intensional part of the data base. The knowledge represented there is a collection of statements used for deduction.

Each of the logical formulae is kept in a certain internal form, corresponding to the way of deduction, described later on. As we have already mentioned, the majority of formulae is of the /i/ form. Every such formula is converted, at the moment of inserting into the data base, to a pair of the following form:

(<conclusion> <premises testing procedure>)

3.2. <u>The knowledge extraction</u>

Because of the menner of storing the knowledge described in the point 3.1, the answer to the question presented to the system does not have to be represented explicite in the data base. The deduction module should be able to obtain all the information included in the data base.

The questions presented to the system are also converted to the logical formulae. Thus, the extraction of knowledge is reduced to the verification of a given formula towards the present content of the data base.

The logical formula representing the question is converted to an appropriate LISP form. Evaluation of such a form is equivalent to examination whether the represented by it formula is true. This form correspond to the normal form of the logical formula /LISP function AND, OR and NOT are used/. The literals are tested by a TESTF function according to the following algorithm:

1. Check the amplifier, trying to find the rule with the conclusion unifiable with the literal under proof. If such a formula does not exist that there is no proof of a given literal;
2. If there is such a formula then:
   a. if it is indicated as an independent formula then STOP with a proof
   b. if it is indicated as a restrictive formula then STOP without a proof;
   c. otherwise evaluate the form associated with the conclusion; if we obtain NIL /false in LISP/ then search the amplifier for another rule and go to 2. If we obtain value different than NIL then STOP

with a proof.
Otherwise Stop without a proof.

It is therefore a so called backward deduction system. The proof goes back from the formula - aim to the facts, applying the formulae from the amplifier in the "Backward" direction.

The answer can be YES or NO or it can be a list of constants depending on the kind of question.

The I order logic has been enriched here with some elements of the II order language. Predicate variables, quantification of these variavles and retrieval of predicates as well as constants have been introduced.

3.3. Access to the data base

The system com municates with the data base through commands of the specially designed language. These commands enable introduction and erasing from the data base.

The basic commands serving the purpose of knowledge extraction are TEST and FIND:

a. TEST A

- looking for the proof of a formula A. Answer YES/NO.

b. FIND $((\lambda\pi_1\ldots\pi_m)(\lambda x_1\ldots x_n)\ A[\pi_1\ldots;x_1\ldots])$

π_i - predicate variables

- retrieval of all the pairs: m-tuple predicates and n-tuple oc constants which satisfy a given formula A.

3.4. Example

The formula presented in the example 1 and a formula below have been introduced into the amlifier.

Sentence:

Wzrost napięcia mięśniówki dwunastnicy może być przyczyną OZT.

/The rise of the tonicity of the duodenum muscular coat may be the reason of acute pancreatitis/

Formula:

```
(IMPLSYM
  (KONJSYM 5((#VAGUE 2) (ACTION X83)
                 (VAG X84 MOC))
   (IMPLSYM
    (KONJSYM 4((#WYDZ-NARZAD 1) (ORGAN X89)
      ((#ORGAN 2) (ORGAN X88) (ONAME X89
                        DWUNASTNICA))
      ((#PART-OF-ORGAN 3) (BODY X87)
          (PNAME X90 MIESNIOWKA)
          (ORGAN X88))
     ((#SICKNESS 4) (SICK X86) (STYPE X91
               FIZJ) (SNAME X92
          NAPIECIE) (BODY X87)))
    ((#RISE 2)(SYMPTOM X85) (SYMPTOM X86)))
  ((#ORGAN 2) (ORGAN X94) (ONAME X95
        TRZUSTKA))
  ((#WYDZ-NARZAD 1) (ORGAN X94))
  ((#SICKNESS 4) (SICK X93) (STYPE X96
        PATO) (SNAME X97 OZT) (BODY X94)))
 ((#IMPLY 3) (INFER X83) (ETIO X85)
          (SICKNESS X93)))
```

Formula corresponding to the question is presented in the Example 2. The amplifier contains the formula describing transitivity of the predicate IMPLY.

Facts - ground literals - were introduced into the nucleus. E.g.

((#BADMEDIC (ALKOHOL)),

(#WYDZ-NARZAD (DWUNASTNICA)), etc.

After converting the formulae of theorems and question into the LISP form its evaluation will find the answer to the question. The answer is of course YES.

4. Conclusion

The results obtained during the work on the system confirmed our direction of research. Our further work will concentrate on constant improvement of the existing modules. At the same time we will undertake attempts of enriching the system with better deductive modules such as resolution in modal logic, default reasoning /Reiter/, FUZZY and Minsky frames.

ACKNOWLEDGEMENTS

The medical text was prepared by a team of physicians from the Postgraduate Education Center in Warsaw under the leadership of Prof. Dr J.Doroszewski. Prof. Doroszewski and his associates have been giving us constant assistance in the interpretation of the medical knowledge included in the presented text. Due to their creative and active cooperation we were able to undertake the elaboration of the described system. We would like to express our cordial gratitude to Prof. Doroszewski and the whole team of doctors.

5. References

Bates, M., The Theory and Practice of Augmented transition Network Grammars in L.Bolc /ed/ Natural Language Communication with Computers

Berry-Rogghe, G.L., Wulz, H., An Overview of PLIDIS a Problem Solving Information System with German as Query Language, in L.Bolc /ed/ Natural Language Question-Answering Systems

Bolc, L., /ed/ Natural Language Communication with Computers, Lecture Notes in Computer Science, Vol 63, Springer -Verlag 1978

Bolc, L. /ed/ Natural Language Based Computer Systems, Hanser Verlag and MacMillan Press, London 1980

Bolc, L. /ed/ Natural Language Question -Answering Systems, Hanser-Verlag and MacMillan Press, London 1980

Bolc, L. /ed/ Representation and Processing of Natural Language, Hanser-Verlag and MacMillan Press, London 1980

Bolc, L., Strzałkowski, T. Transformation of Natural Language into Logical Formulas, Proceedings of the 9Th International Conf. on Comp. Ling. 1982, North Holland Pub. Comp., 1982

Bolc, L., Strzałkowski, T., Natural Language Interface to the Question-Answering System for Physicians, 2nd International Conf. on AI and Information Control Systems of Robots, Conference Proceedings, 1982

Bolc, L., Strzałkowski, T., The Automatic Transformation of Medical Text to a Deductive Data Base, to appear

Bolc, L. /ed/ The Design of Interpreters Compilers, and Editors for Augmented Transition Networks, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983

Codd, F.c., Arnold, R.S., Cadiou, J-M., Chang, C.L., Roussopoulos, N., RENDEZ -VOUS version 1. An Experimental English Language Query Formulation System for Casual Users of Relational Data Base, RJ 2144, IBM Research Lab. San Jose 1978

Burton, R., Brown, J.S., Semantic Grammars: A Technique of constructing Natural Language Interfaces to Industrial Systems, BBN Report No. 3587, Cambridge Ma 1977

Carbonell, J.G., Multi-Strategy Parsing, Dept. of Comp. Sci., Cornegie-Mellon Univ., Pittsburgh Pa, 1981

Chang, C.I., Lee, F.C., Symbolic Logic and Mechanical Theorem Proving, Academic Press, 1973

Dahl, V., Translating Spanish into Logic Through Logic, American Jrnl. of Comp. Linguistics, vol 7., no 3, 1981

Gershman, A.V., Knowledge-Based Parsing, Research Report 156, Yale University, Dept. of Comp. Sci., 1979

Grosz, B., Haas, N., Hendrix, G., Hobbs, J., Martin, P., Moore, R., Robinson, J., Rosenschein, S., DIALOGIC: A Core Natural Language Processing System, Proceedings of the 9th Int. Conf. on Comp. Ling. COLING'82, North Holland, 1982

Klein, N., Implementierung eines Frage- -Antwort-Systems aufder Basis der Predikatenlogik II stufe, Technical Univ. Berlin, 1978

Kochut, K., Towards the Elastic ATN Implementation, in L.Bolc /ed/ The Design of Interpreters, Compilers, and Editors for Augmented Transition Networks, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983

Konrad, E., Formale Semantic von Datenbanksprachen, TUB, Berlin, 1976

Landsbergen, J., Adaptation of Montague Grammar to the Requirements of Parsing, reprint from MC Tract 136, Formal Methods in the Study of Language, J.A.G Groendijk, T.M.V. Jassen, M.B.J. Stokhof /eds/ 1981

Marcus, M.M., ATheory of Syntactic Recognition for Natural Language, The MIT Press, Cambridge Ma, 1980

Martin, W.A., Church, K.N., Patil, R.S., Preliminary Analysis of a Breadth-First Parsing Method, MIT Laboratory Comp. Sci., 1981

Moore, R.C., Problems in Logical Form, Proc. of the 19th Annual Meeting of the ACL, Stanford, California, 1981

Nilsson, N.J., Principles of Artificial Intelligence, Springer-Verlag, Berlin, Heidelberg, New York, 1982

Rosenschein, S.J., Shieber, S.M., Translating English into Logical Form, Proc. of the 20th Ann. Meeting of the ACL, Toronto, 1982

Waltz, D.L., Finin, T.N., Green, F., Conrad, E., Goodman, B., Hadden, G., The PLANES System: Natural Language Access to a Large Data Base, Techn. Rep. T-34, Coordinated Sci. Lab., University of Illinois, 1976

Wilensky, R., Arens, Y., PHRAN - A Knowledge-Based Approach to Natural Language Analysis, Dept. of Comp. Sci., Univ. of California, Berkeley, 1980

Woods, W.A., Transition Network Grammars for Natural Language Analysis, CACM, No 13, 1979, 591-606

Woods, W.A., Kaplan, R.M., Nash-Webber, B., The LUNAR Science Natural Language Information System: Final Report BBN Report No 2378, Bolt Beranek and Newman Inc., Cambridge Ma., 1972

Woods, W.A., Cascaded ATN Grammars, American Jrnl. of Comp. Ling., vol 6 No 1, 1980