

ISOLATING DOMAIN DEPENDENCIES IN NATURAL LANGUAGE INTERFACES

R. Grishman*, L. Hirschman+, and C. Friedman**

*Department of Computer Science, New York University
New York, NY 10012

+Federal and Special Systems Group, Burroughs Corporation
Paoli, PA 19301

**Linguistic String Project, New York University

Abstract

Isolating the domain-dependent information within a large natural language system offers the general advantages of modular design and greatly enhances the portability of the system to new domains. We have explored the problem of isolating the domain dependencies within two large natural language systems, one for generating a tabular data base from text ("information formatting"), the other for retrieving information from a data base. We describe the domain information schema which is used to capture the domain-specific information, and indicate how this information is used throughout the two systems.

Prologue

Computational linguistics is an interesting blend of science and engineering. It is science insofar as we are trying to understand a natural process -- verbal communication. It is engineering insofar as we are trying to manage complexity -- the complexity which, from our present viewpoint, seems inherent in natural language systems.

One tool we have for managing complexity is modular design -- dividing a large system into components of manageable size. This may mean, in particular, separating procedures from knowledge sources and then separating different sources of knowledge. If we "factor" our system in an appropriate way, we may be able to reduce the size not just of individual components but of the system as a whole. Because of our involvement with large natural language systems, we have long been concerned with these issues of modularity [Grishman 1980]. Attacking substantial natural language applications will require that we scale up our already large systems, and we believe that this will be feasible only with systems which have been carefully divided into modules.

One such division is the separation of domain-specific knowledge (sometimes called "world knowledge") from knowledge of the language in general. Such a division not only confers the usual advantages of modularity in facilitating development of a system for a single application, but also greatly enhances the portability of a system to new domains. Portability is especially enhanced if the domain-specific information can be empirically verified and its discovery at least partially automated. What we require is a compact representation of the domain-specific information needed by the components of a natural language system, in a form which can be efficiently utilized by these components. Before we review our own efforts in this direction, a few historical comments are in order on isolating domain-dependent information.

In the early 1970's, the prime concern for most designers was getting these domain-specific ("semantic") constraints into their system; little emphasis was placed on isolating this information from the rest of the system. For example, in the LUNAR system the constraints of operating in a moon rocks world were interwoven with the semantic interpretation procedures [Woods 1972]. Interestingly, one trend of the mid-70's was a tighter integration of domain-specific constraints with general grammatical constraints, in the form of semantic grammars [Burton 1976]. By merging grammatical and 'semantic' constraints, semantic grammars facilitated the construction of small natural language systems. On the other hand, they impeded the capture of grammatical regularities; adding a new syntactic pattern (e.g., reduced relatives) might require adding dozens of productions (one for each allowable combination of semantic classes). They also made it difficult to transport a system to a new domain. As a result, the most recent trend has been towards a careful isolation of domain-specific knowledge (e.g., the RUS System [Bobrow 1980]). In particular, some groups which developed substantial semantic-grammar-based systems, such as

LADDER at SRI [Hendrix 1978] and PLANES at the Univ. of Illinois [Waltz 1978], have now developed syntactic grammars with separate domain information components.

Our systems

For all of the reasons mentioned above -- reduced size and complexity, better capture of grammatical regularities, greater portability, empirical verifiability -- we have been working for the past few years to factor out the domain dependencies from two large natural language systems. One of these is a system for information formatting -- the mapping of natural language text into a tabular data base, for subsequent use in information retrieval and statistical analysis; this system has been used to process radiology reports and hospital discharge summaries [Sager 1978, Hirschman 1982b]. The other is a question-answering system for data retrieval from relational data bases, including in particular those generated by information formatting [Grishman 1978].

In both systems the initial processing -- parsing and transformational decomposition -- is performed by the Linguistic String Parser [Sager 1981]. In formatting, the transformationally regularized parse tree is mapped into an information format; the format is then "normalized" to recover zeroed information and analyze the time structure of the narrative [Hirschman 1981]. For question-answering the operator-operand tree (produced by transformational decomposition) is first translated into a logical form based on first-order predicate calculus; anaphoric phrases are resolved; the logical form is translated into a data base retrieval request; the data is retrieved; and, if necessary, a full-sentence answer is generated incorporating the retrieved data.

The domain information schema (DIS)

The domain-dependent information used by our systems has two basic aspects. First, it characterizes the structure of the information in the domain. Second, it specifies the correspondence between the information structures as they appear in the text and the various internal representations of information in the system.

We call the characterization of the structure of information in the domain a domain information schema or DIS [Grishman 1982]. This characterization consists primarily of a set of semantic classes, the words and phrases which are members of these classes, and the allowable predicate-argument relationships among

these classes in this domain. For example, a schema for a medical domain would contain classes such as PT (patient), VPT (patient-verb), INDIC (indicator of sign or symptom), and BODY-PART:

PT	VPT
patient	experience
pt	complain of
...	...
INDIC	BODY-PART
swelling	abdominal
stiff	neck
...	...

and predicate-argument patterns such as

verb-subject-object:		
VPT	PT	INDIC
complain of	patient	swelling
host-adjective:		
BODY-PART	INDIC	
neck	stiff	

(here class names are given in upper case and members of the classes in lower case).

Certain other properties of these predicates, such as functional dependencies among arguments, are also included in the DIS. For example, in the medical domain there is a functional relationship from tests to patients because each test is of one and only one patient. The DIS is thus similar to data base schemata and to the frame-slot structures of frame-based systems.

Using the DIS

The domain information schema is used most extensively in the parsing stage of the two systems. The predicate-argument constraints of the DIS yield a sublanguage in the sense of Harris [Harris 1968]. These constraints are enforced by a set of selectional restrictions added to the Linguistic String Project English grammar. The task of enforcing these constraints is complicated by the wide variety of surface structures in which a subject-verb-object pattern may appear: declarative, interrogative, and imperative sentences; active and passive voice; in main clauses, relatives, and reduced relatives; etc. The complexity of the restrictions is reduced by the power of the Restriction Language to operate in terms of the string relations (e.g., subject-verb-object or host-modifier relations) [Sager 1975], but it is still substantial. The virtue of this approach, however, is that these restrictions are essentially constant across sublanguages, while the DIS, which will change and grow for new applications, is kept to a minimum.

The following sentence fragment illustrates the use of selectional restrictions to obtain the correct parse:

Blood cultures obtained in the visit to the emergency room prior to admission.

Here the problem is the placement of the prepositional phrase prior to admission, which could modify the directly adjacent noun phrase (the emergency room), but in fact modifies the preceding noun phrase the visit. The selection for prepositional phrases on their hosts is given by the P-NSTGO-HOST list (part of the DIS). The portion of this list relevant for the preposition prior to is:

<u>Prep.</u>	<u>Object of</u>	<u>Host of</u>
	<u>Prep.</u>	<u>Prep. phrase</u>
PREPTIME:	(INDIC:	(INDIC,TEST,VMD,
		VPT,VTR,VSHOW,VHAVE),
	TEST:	(INDIC,TEST,VMD,
		VPT,VTR,VSHOW,VHAVE),
	VMD:	(INDIC,TEST,VMD,
		VPT,VTR,VSHOW,VHAVE),
		...),

This list contains the information that a time preposition (PREPTIME, e.g., prior to) can appear with a VMD (medical action word) as its prepositional object (e.g., prior to admission), with the prepositional phrase modifying another VMD word (e.g., visit); this corresponds to the P-NSTGO-HOST pattern PREPTIME: (VMD: (VMD)). There is no pattern PREPTIME: (VMD: (INST)) which would allow prior to admission to modify the INST word (institution word) emergency room. The application of the selectional constraints ensures that the incorrect parse will be eliminated and the correct one produced.

In order to verify these constraints, the restrictions must determine the semantic class membership of the noun phrases in the sentence. Usually the class of a noun phrase is that of the "core" noun of the phrase. In certain cases, however, the class of the noun phrase as a whole is a function of the classes of both core and adjunct. For instance, a BODY-PART word modified by an INDIC word becomes an INDIC phrase, as in stiff neck. In some cases, the core is "transparent" and the class of the noun phrase is determined by the class of its right or left adjunct, so that, for example, onset of swelling would be in the same class as swelling. To accommodate these situations, the DIS contains rules for such phrasal or "computed" attributes (see the N-LN-COMP-ATT and N-RN-COMP-ATT lists in the appendix). Each time a noun is encountered which can participate in a computed attribute construction, its

adjuncts are checked and, if appropriate, a computed attribute is assigned to be used in further selectional restrictions.

The selectional restrictions serve to exclude many incorrect but syntactically well-formed parses. Constraints on coordinate conjunction (requiring the conjoining of phrases from identical or similar semantic classes), acting together with the computed attribute mechanism, serve to reduce the structural ambiguity of conjoined constructs, always a serious problem [Hirschman 1982a]. The noun phrase anorexia and onset of a stiff neck illustrates this process. There are two possible parses for this phrase, namely (anorexia and onset) of a stiff neck, which is incorrect; and the correct analysis (anorexia) and (onset of a stiff neck). The conjunction mechanism requires that only "similar" elements be conjoined; this rules out the conjunction of anorexia (an INDIC word) and onset (a BEGIN word). However, the phrase stiff neck receives a computed attribute INDIC; and onset is "transparent", so it receives a computed attribute INDIC from its right adjunct stiff neck. Therefore the entire noun phrase onset of a stiff neck has a computed attribute INDIC, and can conjoin (as a phrase) to anorexia, giving the correct parse.

In addition, these selectional patterns can be used to resolve most homographs, that is, to determine the class assignment of words which are members of several classes. For example, the word injection is ambiguous: it can mean inflammation (an INDIC word), as in throat injection, or it can mean shot (a VTR word), as in penicillin injection. The DIS enables a homograph to be disambiguated, provided that sufficient context is present. For example, in the phrase throat injection, the combination INDIC: (BODY-PART) is allowed in the compound noun (N-NPOS) relation, whereas the combination VTR: (BODY-PART) is not allowed (see the appendix for the N-NPOS list). This disambiguation is important because the subsequent mapping into an internal representation (information format or predicate calculus expression) is dependent on the correct identification of the semantic class of each information-carrying word.

The anaphora resolution procedure in the question-answering system relies crucially on the DIS. The same mechanism which uses context to resolve homographs also serves to determine the possible class assignment(s) for an anaphoric phrase. For example, if given the question

Did it show swelling?

the procedure would consult the DIS to determine what classes of subjects can occur with a VSHOW verb (show) and an INDIC object (swelling). The DIS (Appendix, section 2) indicates that the subject in this context can be a BODY-PART or a TEST. The anaphora resolution procedure then searches the current and prior sentences for an antecedent belonging to one of those classes.

The DIS also plays a role in the translation of queries into predicate calculus. Specifically, the information on functional dependencies between arguments of predicates is used in determining the scope of quantifiers and conjunctions. Consider the following two sentences, which have similar syntactic structures:

- (1) How many patients have had an X-ray and a biopsy?
- (2) How many biopsies did patient X and patient Y have?

The first question asks for a single number; in other words, the scope of how many is wider than the scope of and. The second question, however, asks for two numbers: the number of biopsies X had and the number Y had; in this case, the scope of and is wider than that of how many. We know that this is the only possible interpretation of the second question because there are no "group biopsies" -- each biopsy is of one and only one patient. This fact is encoded in the DIS as a functional dependency from TEST to PT (patient) in the triple VHAVE-PT-TEST. By using this functional dependency information, the system is able to assign the correct interpretation to the two questions.

A further application of the DIS (not yet implemented) is the retrieval of "implicit" or omitted information. For example, certain compound noun constructions can be considered to result from the omission of the connector between the two nouns. In these cases, it may be possible to use the verb-subject-object list of the DIS to identify the omitted verb. This can be done by assuming that the head noun of the compound noun phrase will be the subject of the verb, and the modifying noun the object. Thus, given the phrase infectious disease consultant, we have a compound noun whose head is the DOCTOR class, and the modifying noun in the INDIC class. If we search the V-S-O list of the DIS (see appendix) for candidate verbs, we find that a verb of class VTR (treatment) can take a DOCTOR subject and an INDIC object. If, in addition, each class has a distinguished "default" member (e.g., treat for the VTR

class), it may be possible to regularize the compound noun by restoring the omitted information (infectious disease consultant => consultant who treats infectious disease).

Generating Internal Representations

The semantic classes, and the subject-verb-object and host-adjunct patterns are also used to specify the correspondence between the textual and internal representations.

In the current information format for hospital records, most classes map into a unique format column. The formatting procedure records this correspondence as a list of semantic class - format column pairs. For some modifiers, however, such as time modifiers, aspectuals, and negation, the placement of the modifier in the format depends on the class (and hence the placement) of the host; special transformations are provided for the formatting of these modifiers.

The question-answering system has provided slightly greater generality within a two-stage mapping. Syntactically analyzed queries are first mapped into an extended predicate calculus. For each subject-verb-object and host-adjunct pattern in the DIS, we specify a predicate (or set of predicates) and a correspondence between syntactic roles (subject, object, sentence adjunct) and argument positions. Later (after anaphora resolution)

the predicate calculus expression is mapped into a retrieval request on the information format; each predicate is defined as a projection of the information format.

Automated verification and discovery procedures

One of the attractive features of the DIS is that it is empirically verifiable; some of our current research also addresses the possibility of (at least partial) automation of a discovery procedure for portions of the DIS. Semantic classes can be identified within a sublanguage, using techniques of distribution analysis [Hirschman 1975]: for each pair of words in a parsed, regularized sample corpus of a sublanguage, a similarity coefficient is computed based on how many common syntactic environments the two words occurred in (e.g., as the subject of the same verb). "Clusters" of similar words are then formed by grouping together words whose similarity coefficients exceed a certain threshold value. This technique has been used to identify the frequently

occurring members of the major semantic classes of a radiology report domain.

Given the semantic classes, it is then possible to identify the selectional patterns, simply by recording those patterns that occur in (good) parses. This provides verification of the DIS selectional patterns. It also allows collection of data on the relative frequency of occurrence of the various patterns. The frequency data would permit use of a weighting algorithm, in order to "prefer" a parse with more frequently occurring patterns to an alternate parse with less frequently occurring patterns. The "preferential" approach may allow significant enhancement of parsing robustness compared to the "accept/reject" approach currently used. (In the "accept/reject" approach, a parse is either acceptable, or if it violates any constraints, it is rejected; there is no notion of "relative goodness" of several parses). The preferential approach would be particularly useful for incremental development of a DIS in a new sublanguage, where only partial data on selectional patterns is available, and also in highly non-deterministic parsing, such as speech understanding.

One of the issues in automating the discovery procedure for the selectional patterns of the DIS is how to prevent patterns from bad parses from being included in the DIS (and thus allowing even more bad parses). The use of weighted patterns may provide a means for automating the discovery of the DIS, since "correct" patterns are more likely to outnumber random "incorrect" patterns from bad parses. These issues are the subject of an ongoing research project.

Discussion

Both systems described above have been extensively tested. The formatting procedure has been applied to a set of 14 hospital discharge summaries containing over 700 sentences; it is currently being used to process other types of hospital records. The question-answering system has been used on a data base of simplified formatted radiology records. In addition, to test its portability to quite different domains, we have applied the system to a simple data base of student transcripts.*

In the course of this work, we have developed a simple, compact representation of domain-specific knowledge and have thereby substantially reduced the complexity and increased the portability

* The student data base was developed by V. K. Lamson as her master's thesis [Lamson 82].

of our systems. Our experience with the rather different student transcript data base has indicated that not all domain dependencies have yet been isolated, particularly in specifying the mapping from textual to internal representation. Problems arose with the characterization of sentence adjuncts, units of time (semesters instead of days and months), and nouns or noun phrases implying computations (grade point average, enrollment), which we intend to rectify shortly by enriching the DIS.

Our experiments also indicated that relatively limited domain-specific information (primarily a characterization of the structure of information in a domain, rather than specific facts about the domain) can be adequate for certain natural language applications, such as those described. Problems arose more often because the selectional constraints were too "tight" than because constraints deducible from specific facts of the domain were not available. As a result, we are now beginning to experiment with the automatic selective relaxation of these restrictions in order to improve parsing performance.

Acknowledgements

This research was supported in part by National Science Foundation grants MCS 80-02453 from the Division of Mathematical and Computer Sciences and IST 81-15669 from the Division of Information Science and Technology; in part by National Library of Medicine grant 1-R01-LM03933, awarded by the National Institutes of Health, Department of Health and Human Services; and in part by Office of Naval Research contract N00014-75-C-0571, NR 049-347.

References

- [Bobrow 1980] Bobrow, R.J. and Webber, B.L. Knowledge Representation for Syntactic/Semantic Processing, First Annual Nat'l Conf. on Artificial Intelligence, 316-323, AAAI, Stanford, 1980.
- [Burton 1976] Burton, R. Semantic grammar: An engineering technique for constructing natural language understanding systems. BBN Report No. 3453, Bolt, Beranek, and Newman, Cambridge, Mass., 1976.
- [Grishman 1978] Grishman, R., and Hirschman, L. Question Answering from Natural Language Medical Data Bases. Artificial Intelligence 11 (1978), 25-43.

- [Grishman 1980] Grishman, R. Conjunctions and Modularity in Language Analysis Procedures. Proc. COLING 80, 500-503.
- [Grishman 1982] Grishman, R., Hirschman, L., and Friedman, C. Natural Language Interfaces Using Limited Semantic Information. Proc. COLING 82, 89-94 (North-Holland, 1982).
- [Harris 1968] Harris, Z. Mathematical Structures of Language (Interscience, New York, 1968).
- [Hendrix 1978] Hendrix, G., Sacerdoti, E., Sagalowicz, D., and Slocum, J. Developing a natural language interface to complex data. ACM TODS 3, 2 (June 1978), 105-147.
- [Hirschman 1975] Hirschman, L., Grishman, R., and Sager, N. "Grammatically-Based Automatic Word Class Formation," Information Processing and Management Vol. 11 (1975), 39-57.
- [Hirschman 1981] Hirschman, L., and Story, G. Representing Implicit and Explicit Time Relations in Narrative. Proc. IJCAI-81, Vol. 1, 289-295.
- [Hirschman 1982a] Hirschman, L. Constraints on Noun Phrase Conjunction: a Domain-Independent Mechanism. COLING 82 Abstracts, 129-133 (Charles University, Prague, 1982).
- [Hirschman 1982b] Hirschman, L., and Sager, N. Automatic Information Formatting of a Medical Sublanguage. In Sublanguage: Studies of Language in Restricted Semantic Domains (R. Kittredge and J. Lehrberger, eds.), (Walter de Gruyter, Berlin, in press).
- [Lamson 1982] Lamson, V. K. Question-Answering System for an Academic Data Base. Unpublished Master's Thesis, Dept. of Computer Science, New York University, 1982.
- [Sager 1975] Sager, N., and Grishman, R. The restriction language for computer grammars of natural language. Comm. ACM 18, 7 (July 1975), 390-400.
- [Sager 1978] Sager, N. Natural Language Information Formatting: The Automatic Conversion of Texts to a Structured Data Base. In Advances in Computers 17 (M.C. Yovits, ed.), 89-162 (Academic Press, NY, 1978).
- [Sager 1981] Sager, N. Natural Language Information Processing (Addison-Wesley, 1981).
- [Waltz 1978] Waltz, D. An English language question answering system for a large relational data base. Comm. ACM 21, 7 (July 1978), 526-539.
- [Woods 1972] Woods, W. A., Kaplan, R. M., and Nash-Webber, B. The lunar sciences natural language information system: Final report. Report 2378, Bolt, Beranek, and Newman, Cambridge, Mass., 1972.

APPENDIX

AN EDITED DOMAIN INFORMATION SCHEMA FOR A MEDICAL SUBLANGUAGE

1. SUBLANGUAGE SEMANTIC CLASSES

- * Below are some of the sublanguage classes used in the medical domain information schema; note that classes may contain words from different syntactic classes. The 15 classes shown below were selected for illustrative purposes from the over 50 classes in the full DIS.
- * Classes are given in the format: (abbreviated) CLASS NAME, [explanation of name], followed by a few class members.

PT	DOCTOR	INST	TEST	RX
{patient}	{doctor}	{medical institution}	{test}	{medication}
patient	doctor	hospital	x-ray	penicillin
pt	consultant	emergency room	red cell count	hydration
..
INDIC	BODY-PART	AMT	PREPTIME	BEGIN
{indicator of sign/symptom}	{part of body}	{amount}	{time preposition}	{beginning}
swelling	neck	severe	during	onset
stiff	throat	high	after	start
disease	muscle	low	prior to	beginning
injection	abdominal
[=inflammation]	...			
..				

VMD [medical action]	VPT [patient verb]	VTR [treatment]	VSHOW [show]	VHAVE [possession, association]
admission	complain of	treat/ment	show	have
discharge	experience	inject/ion	reveal	
visit	suffer from	prescribe	indicate	
..	

2. ALLOWABLE PREDICATE-ARGUMENT RELATIONSHIPS

- * The following lists are edited versions of the various patterns
- * of selectional relations stated in terms of the (medical) domain
- * semantic classes.

LIST V-S-O =

- * Verb-Subject-Object allowable patterns, given in the form:
- * VERB: (SUBJECT1: (OBJECT11,...OBJECT1n),
- * SUBJECT2: (OBJECT21,...OBJECT2m),
- * ...
- * Thus a VPT verb (e.g., complain of) can occur with a
- * PT subject and INDIC object (OK: patient complained of fever),
- * but not with an INDIC subject and PATIENT object
- * (NO: fever complained of patient).
- * VMD: (DOCTOR: (PT)),
- * VPT: (PT: (INDIC)),
- * VTR: (DOCTOR: (INDIC, PT, RX, TEST, VTR),
- * INST: (INDIC, PT, RX, TEST, VTR)),
- * VSHOW: (BODY-PART: (INDIC),
- * TEST: (INDIC)),
- * VHAVE: (PT: (DOCTOR, INDIC, TEST, VTR),
- * BODY-PART: (INDIC)),
- * ...

LIST N-NPOS =

- * Noun-compoundNoun list
- * describes which classes of head noun can be modified by
- * which classes of compound noun (NPOS) modifier in the form:
- * HEAD-NOUN1: (MODIFIER-NOUN11,...,MODIFIER-NOUN1n),
- * HEAD-NOUN2: (MODIFIER-NOUN21,...,MODIFIER-NOUN2m),
- * ...
- * Thus the compound noun INDIC :(BODY-PART), as in
- * throat injection, is allowable, but the compound noun
- * BODY-PART :(INDIC), as in injection throat, is not.
- * DOCTOR: (BODY-PART [ent consultant], INDIC),
- * INDIC: (BODY-PART, INDIC),
- * INST: (INST [hospital emergency room]),
- * PT: (INDIC),
- * VMD: (INST, INDIC),
- * VTR: (INST, RX, INDIC),
- * ...

LIST N-ADJ =

- * Noun-Adjective list
- * describes which classes of head noun can be modified by
- * which classes of adjectives in the form:
- * HEAD-NOUN1: (ADJECTIVE11,...,ADJECTIVE1n),
- * HEAD-NOUN2: (ADJECTIVE21,...,ADJECTIVE2m),
- * ...
- * Thus the adjective-noun combination given by BODY-PART :(INDIC),
- * as in stiff neck is allowed, but BODY-PART: (AMT) is not
- * (NO: severe neck).
- * BODY-PART: (BODY-PART, INDIC),
- * INDIC: (AMT, BODY-PART, INDIC),
- * PT: (INDIC),
- * RX: (VTR [prophylactic penicillin],
- * BODY-PART [cardiac medication]),
- * TEST: (AMT, BODY-PART),
- * ...

LIST P-NSTGO-HOST =

* Preposition-NounObject-Host
 * describes which prepositional phrases can modify which hosts.
 * These patterns have the form:
 * PREPOSITION1: (NOUN-OBJECT11: (HOST111, HOST112,...),
 * NOUN-OBJECT12: (HOST121, HOST122,...),
 * ...),
 * PREPOSITION2: (NOUN-OBJECT21: (HOST211, HOST212,...),
 * ...),
 * ...
 * Thus the prepositional phrase given by PREPTIME:(VMD:(INDIC)),
 * as in swelling after admission is allowable, but there
 * is no pattern PREPTIME:(BODY-PART:(INDIC)), e.g., no phrase
 * swelling after neck.
 * 'IN': (BODY-PART: (INDIC, TEST),
 * INST: (DOCTOR, PT),
 * ...),
 * PREPTIME: (INDIC: (INDIC, TEST, VMD, VPT, VTR, VSHOW, VHAVE),
 * TEST: (INDIC, TEST, VMD, VPT, VTR, VSHOW, VHAVE),
 * VMD: (INDIC, TEST, VMD, VPT, VTR, VSHOW, VHAVE),
 * ...),
 * ...

3. COMPUTED ATTRIBUTE LISTS

* There are two computed attribute lists, one for head noun + left adjuncts,
 * and one for head noun + right adjuncts. For computed attributes,
 * the list must specify the class of the head, the "computed" class,
 * and the class of the adjunct. The lists have the form:
 * HEAD-NOUN1: (COMPUTED-CLASS11: (ADJUNCT111, ADJUNCT112,...),
 * COMPUTED-CLASS12: (ADJUNCT121, ADJUNCT122,...),
 * ...),
 * HEAD-NOUN2: (COMPUTED-CLASS21: (ADJUNCT211, ADJUNCT212,...),
 * ...),
 * ...
 * Thus a BODY-PART head noun will give a "computed attribute" INDIC
 * when modified by an INDIC left modifier (stiff neck); a BEGIN
 * head noun will give an INDIC computed attribute when modified by
 * either a left or a right INDIC modifier: fever onset or
 * onset of fever.

N-LN-COMP-ATT =

* Noun = Computed Attribute with LeftNoun adjunct
 * BODY-PART: (INDIC: (INDIC)),
 * BEGIN: (INDIC: (INDIC), TEST: (TEST), RX: (RX),
 * VMD: (VMD), VTR: (VTR)),
 * ...

N-RN-COMP-ATT =

* Noun = Computed Attribute with RightNoun adjunct
 * BEGIN: (INDIC: (INDIC), TEST: (TEST), RX: (RX),
 * VMD: (VMD), VTR: (VTR)),
 * ...