

33

Strategies for Understanding Structured English

Alain Bonnet

Psychological work on memory, in particular by Bartlett (1932), has led to the conclusion that people faced with a new situation use large amounts of highly structured knowledge acquired from previous experience. Bartlett used the word *schema* to refer to this phenomenon. Minsky (1975), in his famous paper, proposed the notion of a *frame* as a fundamental structure used in natural language understanding, as well as in scene analysis. I will use the former term in the rest of this chapter, in spite of its general connotation.

The main thesis defended by Bartlett was that the phenomena of memorization and remembering are both constructive and selective. The hypothesis has more recently been revived by psychologists working on discourse structure (Collins, 1978; Bransford and Franks, 1971; Kintsch, 1976). Various experiments performed on subjects who were told stories and then asked to describe what they remembered showed that people not only forget facts but add some. Moreover, they are unable to distinguish between what they have actually heard and what they have inferred. People hearing a story make assumptions, which they might revise or refine as more information comes in, either confirmatory or contradictory. Making such assumptions entails building (or retrieving) models of the expected text contents. A corollary of this process is that if the story adequately fits the model people have in mind, the story will be understood more easily.

Although it is difficult to give a formal definition of what constitutes a coherent text, it is an accepted notion that sentences that comprise it are

This chapter is based on a technical memo (HPP-79-25) from the Heuristic Programming Project, Department of Computer Science, Stanford University. Used with permission.

linked by cause-effect relationships, chronological orderings, and the like. Flashbacks are not contradictory with coherence, but they can make the text more difficult to comprehend. Texts dealing with specific domains seem to be structured in terms of topic. Consequently, an important problem to face is recognizing the different topics and deciding when a *shift in topic* occurs.

Several frame-based languages, such as KRL (Bobrow and Winograd, 1977), the “units package” (Stefik, 1979), and FRL (Roberts and Goldstein, 1977), implement the basic concepts underlying frames, or schemata. A schema contains slots. They can be viewed as variables that will be bound to data. Each slot contains “facets” (FRL), “aspects” (units package), or “descriptors” (KRL), which specify how to fill the slots, for example, specifying the type of values acceptable (numeric, strings of characters), the range of possible values, values to assign by default, or attached procedures describing what to do if the slot is filled in (this is a way to make inferences). Slots may be organized into hierarchical schemata, in which case values may be inherited from one schema to a more specialized one. This hierarchy and concomitant inheritance avoids any duplication of common properties.

BAOBAB is an experimental natural language interface to the MYCIN system and is based on the idea that clusters of topics can be represented in framelike structures. Its design has been inspired by observations regarding text coherence, shifts in topic, temporal orderings, etc. Figure 33-1 shows a short dialogue between a physician and the program. The English that is generated is somewhat stilted, but its main purpose is to be nonambiguous. In this chapter are presented the principles and methods that permitted the implementation of a program with these capabilities.

33.1 Related Work and Goals

AI researchers have explored a variety of strategies to recognize shifts in topic occurring during dialogues or written texts. One issue faced is the necessity to narrow the space of possible referents of a linguistic object by focusing at different levels of detail. Thus Grosz (1977) studied the role of focus in the interpretation of utterances and its relation to domain structure. She used the task structure to resolve definite noun phrases in task-oriented dialogues. Sidner (1979) extended this work to determine the use of focusing in the resolution of pronoun references and other kinds of anaphora occurring in dialogues. Rosenberg (1977) has studied how themes are linked via references in newspaper articles.

Another major reason for focusing is to avoid combinatorial explosion. The fear of such an explosion is mainly motivated by the large number of inferences that may be made if all possible frames are activated. In fact,

I am ready

** THIS IS A 26 YEAR OLD MALE PATIENT

My understanding is:
The age of the patient is 26
The sex of the patient is male

** FIVE DAYS AGO, HE HAD RESPIRATORY-TRACT SYMPTOMS

What is his name?

** JO

My understanding is:
The name of the patient is Jo
Respiratory-tract is one of the symptoms that the patient had

** A COUPLE OF DAYS BEFORE THE ADMISSION, HE HAD A MALAISE

Please give me the date of admission

** MARCH 12, 1979

My understanding is:
The patient was admitted at the hospital 3 days ago
Malaise is one of the symptoms that the patient had 5 days ago

FIGURE 33-1 Short sample dialogue. The physician's inputs appear in capital letters after the double asterisks.

some of the frames might rule out others, thus enabling the space of possible inferences to be pruned. This issue has also been raised by Charniak (1978). Embodying world knowledge in frames (Minsky, 1975) or scripts (Abelson, 1973; Schank and Abelson, 1975) led to the development of programs that achieved a reasonably deep level of understanding, for example, GUS (Bobrow et al., 1977), NUDGE (Goldstein and Roberts, 1977), FRUMP (DeJong, 1977) and SAM (Cullingford, 1977).

BAOBAB and the other programs mentioned so far have a common feature: they do not interpret sentences in isolation. Rather, they interpret in the context of an ongoing discourse and, hence, use discourse structure. BAOBAB also explores issues of (a) what constitutes a model for structured texts and (b) how and when topic shifts occur. However, BAOBAB is interested neither in inferring implicit facts that might have occurred temporally between facts explicitly described in a text nor in explaining intentions of characters in stories (main emphases of works using scripts or plans). Our program focuses instead on coherence of texts, which is mainly a task of detecting anomalies, asking the user to clarify vague pieces of information or disappointed expectations, and suggesting omissions. The domain of application is patient medical summaries, a kind of text for which language-processing research has mainly consisted of filling in formatted grids without demanding any interactive behavior (Sager, 1978). BAOBAB's objectives are to understand a summary typed in "natural med-

ical jargon” by a physician and to interact by asking questions or displaying what it has understood.

The program uses a model of the typical structure of medical summaries, which consists of a set of related schemata, described below. BAOBAB uses both its medical knowledge and its model of the usual description of a medical case to interpret the dialogue or the text and to produce an internal structure usable by MYCIN. The program then uses this information to guide a standard consultation session.

BAOBAB behaves like a clerk or a medical assistant who knows what a physician has to describe and how a malady is ordinarily presented. It reacts to violations of the model, such as a description that ignores symptoms or that fails to mention results of cultures that have been drawn. It does not attempt to use its knowledge to infer any diagnosis but, in certain cases, can draw inferences that will facilitate MYCIN's task. BAOBAB uses these capabilities to establish relationships between the concepts stated. This facilitates interpretations of what is said. For example, BAOBAB knows that “semi-coma” refers to the state of consciousness of the patient and “hyperthyroidism” to a diagnosis. One use of the program would be to allow the physician to volunteer information before or during the consultation. This feature would respond to the common frustration expressed by some users who object to having to wait for MYCIN to ask a key question before they can tell it about a crucial symptom.

BAOBAB consists of (a) a parser that maps the surface input into an internal representation, (b) a set of schemata that provide a model of the kind of information that the program is ready to accept and of the range of inferences that it will be able to draw, (c) episode-recognition strategies that allow appropriate focusing on particular pieces of the texts, and (d) an English-text generator used to display in a nonambiguous fashion what has been understood. As described in Chapter 5, this generator was already available in MYCIN. The main emphasis here will therefore be on the description of schemata and schema-activation strategies. These techniques have been successfully implemented, using Interlisp (Teitelman, 1978), in a program connected with MYCIN's data base and running on the SUMEX computer at Stanford.

33.2 Schemata and Their Relations

Medical summaries can be viewed as sequences of episodes that correspond to phrases, sentences, or groups of sentences dealing with a single topic. Each such topic may be represented by a schema. Processing and understanding a text consist of mapping episodes in the text onto the schemata that constitute the model. Matching a schema can be discontinuous; that is, two episodes referring to the same schema need not necessarily be jux-

taped (they might be separated by an episode referring to another schema). We will refer to this phenomenon as a *temporary schema-shift*.

A typical scenario is as follows. The medical case is introduced with general information, such as the date and the reason for admission to the hospital. Then the patient is presented (name, age, . . .). Symptoms (noted by the patient) and signs (observed by the physician) are described. A physical exam is usually performed, and cultures are taken for which results are pending or available. The structure of such a text can be captured in a sequence of schemata, one of which is shown in Figure 33-2. These texts are usually well structured. Redundancies can appear, but discrepancies are rather rare (although they must be detected when they occur). Expectations are usually satisfied.

A typical BAOBAB schema contains domain-specific knowledge and resembles a frame (Minsky, 1975) or script (Schank and Abelson, 1975) or unit (Stefik, 1979). Relevant slots define expected values, default values, and attached procedures. Attributes relating to the same topic are gathered into these schemata. There is some overlap between them (such as WEIGHT, which can occur in the identification of the patient as well as in the results of a physical exam). Each schema contains two types of slots: global slots (comments, creation date, author's name, how to recognize the schema, what is the preferred position of the schema within summaries) and individual slots (which correspond to MYCIN's clinical parameters). Each individual slot contains *facets* specifying how to fill it in or what actions to take when it has been filled in (by procedural attachment).

Global slots are mainly used to decide whether a part of the text being analyzed suggests or confirms a schema or how the confirmation of one schema causes another one to be abandoned. The slots CONFIRMED-BY and SUGGESTED-BY point to lists of slots belonging to the schema. The first defines the schema (characteristic slots), whereas the other is nonessential for confirming the schema. The slots TERMINATED-BY and PREF-FOLLOWED-BY specify relationships of mutual exclusion and partial ordering between schemata. All these slots are described in more detail in the section devoted to strategies for activating schemata. Nonglobal slots are always attributes grouped within a schema. Each is, in turn, a schema whose slots are the facets mentioned above (Roberts and Goldstein, 1977).

33.2.1 An Example of a Schema

In the \$DESCRIPT schema (Figure 33-2), the first three global slots (AUTHOR, CREATION-DATE, and COMMENT) are used for documentation, whereas the next four are used to define strategies for schema-shifts (see below). Then six individual slots (corresponding to parameter names) define the schema. Each of them is described by subslots, or facets, some of which (e.g., EXPECT, TRANS, LEGALVALS, CHECK, PROMPT) already exist in the structure of MYCIN's knowledge base. Others have been

\$DESCRPT

AUTHOR: BONNET
 CREATION-DATE: OCT-10-78
 COMMENT: Patient identification
 CONFIRMED-BY: (NAME AGE SEX RACE)
 TERMINATED-BY: (\$SYMPTOM)
 SUGGESTED-BY: (WEIGHT HEIGHT)
 PREF-FOLLOWED-BY: (\$SYMPTOM)

NAME

EXPECT: ANY
 TRANS: ("the name of" *)
 TOBEFILLED: T
 WHENFILLED: DEMONNAME

AGE

EXPECT: POSNUMB
 TRANS: ("the age of" *)
 CHECK: (CHECK VALU 0 100.0 (LIST "Is the patient really"
 VALU "years old?") T)
 TOBEFILLED: T
 WHENFILLED: SETSTATURE

SEX

EXPECT: (MALE FEMALE)
 TRANS: ("the sex of" *)
 TOBEFILLED: T
 WHENFILLED: SEXDEMON

RACE

EXPECT: (CAUCASIAN BLACK ASIAN INDIAN LATINO OTHER)
 TRANS: ("the race of" *)

WEIGHT

EXPECT: POSNUMB
 TRANS: ("the weight of" *)
 CHECK: (CHECK VALU LIGHT HEAVY (LIST "Does the patient
 really weigh" VALU "kilograms?") T)

HEIGHT

EXPECT: POSNUMB
 CHECK: (CHECK VALU SMALL TALL (LIST "Is the patient
 really" VALU "centimeters tall?") T)

FIGURE 33-2 Schema of a patient description.

created to allow the program to intervene during the course of the dialogue. For example, when the slot TOBEFILLED holds the value T (true), it means that the value of the variable must be asked if the physician does not provide it. The WHENFILLED feature specifies a procedure to run as soon as the slot is filled in. This is the classic way of making inferences. For example, SETSTATURE specifies narrower ranges of weight and height for a patient according to his or her age.

33.2.2 Facets

Expected and legal values. EXPECT is used for single-valued parameters, whereas LEGALVALS is used for multi-valued parameters (see Chapter 5). They both give a list of possible values for an attribute.

Linguistic information. TRANS always contains a phrase in English describing the parameter; it is used for generating translations of rules and other semantic entities. PROMPT contains a question, in English, that asks the user about the corresponding parameter. It is used, in addition to the usual way MYCIN asks for information, to clarify a concept recognized as “fuzzy.” For example, entry of the clause “THE PATIENT DRINKS 6 CANS OF BEER EVERY MORNING” leads BAOBAB to ask “Is the patient alcoholic?” since MYCIN has no explicit knowledge about alcoholic beverages, but can recognize such keywords as drink or alcohol. CHECK contains a question that can be used to request verification whenever a value outside the normal range has been given.

TOBEFILLED. If the TOBEFILLED facet of an attribute is set to T (true), it means that the slot has to be filled. Concretely, this means that if the slot has not yet been filled when the schema is abandoned, the attached request will be carried out. This does not necessarily mean that the parameter is essential from a clinical point of view; it may be essential for communication purposes.

33.2.3 Procedural Attachment

In BAOBAB, there are two kinds of procedural attachment. The first, called WHENFILLED, allows associated actions to be carried out depending on conditions local to the slot. It is analogous to the “demons” of Selfridge (1959) or Charniak (1972). The second kind of attachment, called PREDICATE, is used to specify how to fill a slot and is mentioned last. These facets allow BAOBAB to:

- a. *Produce inferences.* If the attribute of a clause that has just been built has an attached procedure, it can trigger the building of another clause; for example, INFERFEVER is run as soon as the temperature is known and can lead to a clause such as “The patient is not febrile.”
- b. *Narrow a range of expected values.* Consider, for example, the weight of a patient. This has *a priori* limits, by default, of 0 and 120 kilograms. This range is narrowed according to the age of the patient as soon as the latter is known.
- c. *Make predictions.* An event like “a lumbar puncture” can cause predictions about “CSF data” (not about their values, but about the fact that

they will be mentioned). These predictions will be checked, and appropriate questions will be asked if they remain unfulfilled as the dialogue proceeds.

- d. *Dynamically modify the grammar.* A semantic category like <PATIENT> can be updated by the name of the patient as soon as it is known. This update is done by the procedure DEMONNAME as indicated in Figure 33-2.
- e. *Specify how to fill a slot.* Sometimes a procedure expresses the most convenient way to match a category. This kind of procedure has been called a “servant.” For example, the best way to match a <VALUE> is to know that it points to its corresponding <ATTRIBUTE>. This is much simpler than examining the list of 500 values in the dictionary.

33.2.4 Default Values

BAOBAB distinguishes among three kinds of default values:

- a. Some parameters have default values that are negations of symptoms; for example, TEMPERATURE has “98.6 F” as a default value (negation of fever), and STATE-OF-CONSCIOUSNESS has “alert” as a default value (negation of altered consciousness).
- b. Other parameters depend on the result of a medical exam or procedure, and in such cases the default value is simply “unknown.” Pointing out an unknown value to the physician might remind him or her that the procedure has in fact been carried out and that a result should have been mentioned. An example of such a default value is that for the parameter STATE-OF-CHEST, which depends on an x-ray.
- c. Finally, some parameters inherit a value from another variable; for example, the date of a culture might reasonably be the date of admission to the hospital (if the infection is not hospital-acquired).

Note that any default value assumed by the program is explicitly stated. This feature allows the user to override the default value when in disagreement with it (a mandatory feature because a default value might be used later by the consultation program and therefore be taken into account in the formation of the diagnosis).

33.3 The Grammar

In a technical domain, where specialists write for specialists, terseness of style is widespread (e.g., “T 101.4 rectal”). Thus a syntactic parsing does not provide enough additional information to justify its use for text com-

prehension. Instead, a computer program can use a semantically oriented grammar. This grammar makes the parsing process unambiguous and therefore efficient. Discussions of this point can be found in Burton (1976) and Hendrix (1976).

BAOBAB's parser uses a context-free augmented grammar [cf. the augmented transition network of Woods (1970)]. A grammar rule specifies (1) the syntax, (2) a semantic verification of the parsed tree resulting from the syntactic component, and (3) a response expression used to build one or several clauses. The grammar is divided into specific and nonspecific rules.

Specific grammar rules are associated with the slots of schemata and describe the way these can be mentioned at the surface level. Categories used in the rules are things such as <PATIENT>, <SIGN>, and <DIAGNOSIS>. This link between the grammar and the schemata provides a means to try, by priority, those grammar rules that are appropriate to the schema under consideration. Furthermore, it provides a means to postpone the risk of combinatorial explosion due to the large number of grammar rules (due to the specificity of the categories used in the productions).

Nonspecific grammar rules use general concepts such as <ATTRIBUTE>, <OBJECT>, and <VALUE>, which are commonly used to represent knowledge in systems. This kind of rule is general enough to be used in other domains; but once the syntax has been recognized, these rules must undergo a semantic check in order to verify that, say, values and attributes fit together, hence the importance of the augmentation of the grammar mentioned above.

Specific grammar rules enable the system to recognize peculiar constructs. For example, "120/98" and "98 F" do not belong to well-known syntactic classes but have to be recognized as values for blood pressure and temperature. Grammar rules such as

<VITAL> → <BP> <HIGH/LOW>

<VITAL> → <TEMP><TEMPNUM>|<TEMP><NUM>(DEGREES)

are used to parse "BP 130/94" or "T 98 F." The category <TEMPNUM> has an attached procedure, a specific piece of code that recognizes "F" as Fahrenheit, detaches it from "98," verifies that 98 is a reasonable value for a temperature, and finally returns "98 degrees" as the value of the temperature.

The following are examples of the "syntax" of purely semantic rules:

<sentence> → <patient> <experience> <symptom> <time>

<symptom> → <modifier> <symptom>

<patient> → patient | <name>

- <name> → (the name of the patient, usually encountered at the beginning of the text)
- <experience> → complain of | experience | <have>
- <symptom> → headache | malaise | chill | . . .
- <modifier> → severe | painful | . . .
- <have> → has | had | . . .
- <time> → <num> <time-unit> ago | on <date>
- <time-unit> → day | week | . . .
- <num> → 1|2|3| . . .
- <date> → a date recognized by an associated LISP function

This subset of the grammar enables the program to recognize inputs such as the following:

1. NAPOLEON COMPLAINED OF SEVERE HEADACHE 3 DAYS AGO
2. BILL EXPERIENCED MALAISE ON SEPT-22-1978
3. JANE HAD CHILLS ON 10/10/78

Examples of purely syntactic rules are as follows:

- <SENTENCE> → <NP> <VP>
- <NP> → <NOUN> | <ADJ> <NOUN> | <DET> <ADJ>
<NOUN> | <DET> <NOUN> | . . .
- <VP> → <VERB> | <VERB> <NP> | <VERB> <PREPP>
- <PREPP> → <PREP> <NP>

where <NP> stands for noun phrase, <VP> for verb phrase, <DET> for determiner, <PREPP> for prepositional phrase and <PREP> for preposition. The set of rules enables the system to recognize input sentence 1 above (except for the notion of time), as shown in the syntactic tree of Figure 33-3.

When the semantic component interprets such a syntactic tree, it checks that <NOUN> is matched by a person (whereas the direct use of <PATIENT> would make useless such a verification). Input sentences such as the following would thus be rejected:

4. THE BOAT COMPLAINED OF HEADACHE
5. BILL COMPLAINED OF A SEVERE LEG

Numerous systems use a representation based on the notion of object-

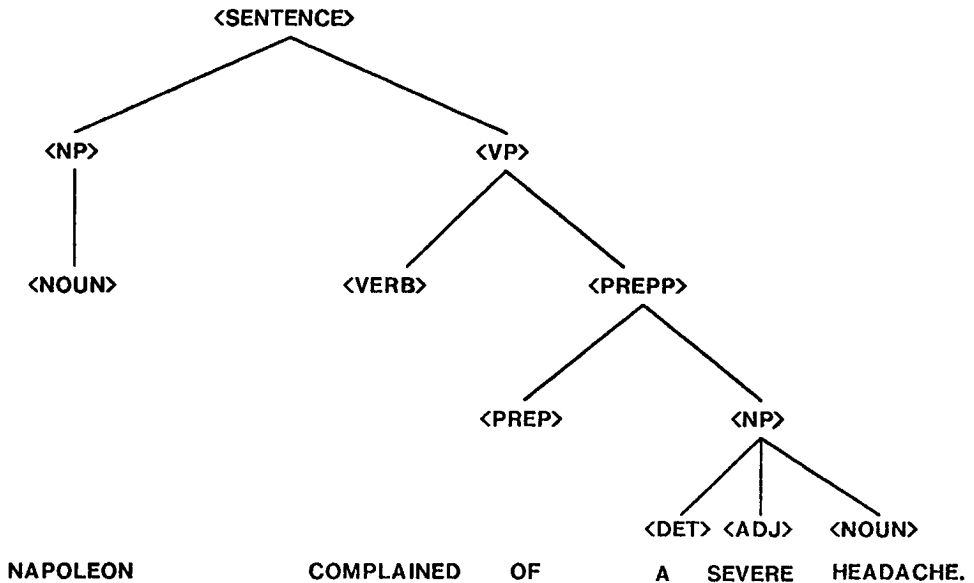


FIGURE 33-3 A conventional syntactic tree.

attribute-value triples with an optional associated predicate-function. In such domains, one can define grammar rules such as:

$\langle \text{SENTENCE} \rangle \rightarrow \langle \text{OBJECT/ATTRIBUTE} \rangle \langle \text{PREDICATE-FUNCTION} \rangle \langle \text{VALUE} \rangle$

$\langle \text{OBJECT/ATTRIBUTE} \rangle \rightarrow \langle \text{ATTRIBUTE} \rangle \text{ OF } \langle \text{OBJECT} \rangle \mid \langle \text{OBJECT} \rangle \langle \text{ATTRIBUTE} \rangle$

$\langle \text{OBJECT} \rangle \rightarrow \text{PATIENT} \mid \text{CULTURE} \mid \text{ORGANISM} \mid \dots$

$\langle \text{ATTRIBUTE} \rangle \rightarrow \text{ISATTRIBUTE}$ (attached procedure specifying how to recognize an attribute)

$\langle \text{PREDICATE-FUNCTION} \rangle \rightarrow \langle \text{SAME} \rangle \mid \langle \text{NOTSAME} \rangle \mid \dots$

$\langle \text{SAME} \rangle \rightarrow \text{IS} \mid \text{HAS} \mid \dots$

$\langle \text{VALUE} \rangle \rightarrow \text{ISVALUE}$ (attached procedure specifying how to recognize the value of an attribute)

Such “syntactico”-semantic rules enable the recognition of input sentences such as:

6. THE TEMPERATURE OF THE PATIENT IS 99

7. THE MORPHOLOGY OF THE ORGANISM IS ROD

The complete form of the <SENTENCE> rule is displayed below. The first line is the syntax, the second is the augmentation, and the third is the response. CHECKAV (check attribute value) is a function of two arguments, <ATTRIBUTE> and <VALUE>, that returns “true” if the value matches the attribute, in which case the response expression is produced; otherwise, the semantic interpretation has failed.

```
((<OBJECT/ATTRIBUTE> <PREDICATE-FUNCTION> <VALUE>)
((CHECKAV <ATTRIBUTE> <VALUE>)
(LIST <PREDICATE-FUNCTION> <ATTRIBUTE> <VALUE>)))
```

It is interesting to note that the predicate function is usually a verb phrase, and the <ATTRIBUTE> OF <OBJECT> sequence a noun phrase, as is <VALUE>. This means that a syntactic structure is being implicitly used.

The interpreter progresses in a left-to-right and top-down fashion, with backtracking. Whenever a grammar rule is satisfied but a part of the input remains to be analyzed, the remaining part is given back to the control structure, which then can invoke special processes; for example, a conjunction at the head of the remaining input can trigger an attempt to resolve it as an elliptical input. Thus in “ENGLISH PEOPLE LOVE BLONDS AND DRINK TEA,” the second part can be analyzed as “English people drink tea.” The algorithm implemented for handling elliptical inputs has been inspired by LIFER (Hendrix, 1976). When an input fails to be recognized, the interpreter assumes that a part of the input is missing or implicit, and it looks at the preceding utterance. If parts of the input match categories used in the grammar rule satisfied by the earlier input, it then assumes that the parts that have no correspondence in the new input can be repeated.

33.4 Schema-Shift Strategies

A language describing choices between schemata, and therefore schema-shift strategies, should include an attempt to answer the following questions: How is a schema focused, confirmed, or abandoned? What are the links between schemata (such as exclusive or sequencing relations)?

33.4.1 Suggest vs. Confirm

Bullwinkle makes the distinction [Bullwinkle (1977); see also Sidner (1979)] between potential and actual shifts of focus, pointing out that the cues

suggesting a new frame must be confirmed by a subsequent statement in order to avoid making unnecessary shifts. This phenomenon is handled in a different fashion in BAOBAB. Instead of waiting for the suggestion to be confirmed, a qualitative distinction is made between the slots of a frame. The ones marked as suggesting but not confirming are regarded as weak clues and will not lead to a shift of focus, whereas the ones marked as confirming (hence suggesting) are sufficiently strong clues to command the shift. This distinction can be illustrated by the following two examples:

1. "The patient was found comatose. She was admitted to the hospital. A lumbar puncture was performed. She denied syncope or diplopia . . ."
2. "The patient was found comatose. He was admitted to the hospital. The protein from CSF was 58 mg% . . ." (CSF = cerebrospinal fluid)

In Example 1, the lumbar puncture suggests CSF results that are not given (weak clue). In Example 2, a detail of CSF results (strong clue) is given directly ("the protein"). In other words, the physician jumps into detail, and the frame is directly confirmed.

33.4.2 Top-down vs. Bottom-up

Sometimes the schema is explicitly announced, as in "results of the culture." This is a name-driven invocation of the schema. More often, the instantiation of the schema is content-driven. The clues used are the attributes associated with the schema, their expected values (if any), and other concepts that might suggest the frame. For example, "skin" is related to "rash," which belongs to the physical exam frame. These are indeed very simple indices. Research on more sophisticated methods for recognizing the relevant schema, such as discrimination nets, have been suggested (Charniak, 1978).

33.4.3 Termination Conditions

A simple case in which a schema can be terminated is when all of its slots have been filled. This is an ideal situation, but it does not occur very often. Another case is when the intervention of a schema implies that another schema is out of focus, which could be, but is not necessarily, the result of chronological succession. In general, this phenomenon occurs when the speaker actually starts the plot after setting the characters of the story. There is no standard way to decide when the setting is finished. However, as soon as the story actually starts, the setting could be closed and possibly completed with default values or with the answers to questions about whatever was not clear or omitted. A TERMINATED-BY slot has been created

to define which schemata can explicitly terminate others; for example, the \$SYMPTOM schema usually closes the \$DESCRIPT schema (name, age, sex, race), as it is very unlikely that the speaker will give the sex of the patient in the middle of the description of the symptoms. This fact is due to the highly constrained nature of the domain.

33.4.4 Termination Actions

When a schema is terminated, the program infers all the default values of the unfilled slots. It also checks whether the expectations set during the story have been fulfilled. These actions can be performed only when a shift has been detected or at the end of the dialogue; otherwise, the program might ask too early about information that the user will give later. In the case where a schema has been exhausted (all its slots filled), an *a priori* choice with regard to the predicted next schema is made. This choice is possible by using a PREFERABLY-FOLLOWED-BY pointer that, in the absence of a bottom-up (data-driven) trigger for the next schema, decides in a top-down fashion which schema is the most probable to follow at a given point.

33.4.5 Schema-Grammar Links

Specific grammar rules described earlier are always associated with clinical parameters and therefore with schemata. This link is interesting from two points of view:

- a. The interpreter takes advantage of this relationship to try specific rules in order of decreasing probability of relevance to the schema currently in focus. There is no quantitative notion of probability, but the preferred sequencing causes the trial according to priority not only of grammar rules associated with the activated schema, but also of the ones of the preferred successor, in case an unforeseen shift occurs. Rules are reordered whenever a schema-shift occurs, which explains why the more disorganized presentations of a text take longer to be parsed.
- b. The parser can examine the content of a schema during the semantic interpretation of an input. For example, it can check the correspondence of an attribute and a value. It can also trigger a question whose answer is needed to interpret the current input. Therefore, there is a two-way connection between schemata and the grammar. This link is one of the key ideas underlying the interactive behavior of the program.

33.4.6 Comparison with Story-Grammars

Other methods have been proposed to take advantage of the coherent structure of texts. Psychologists and linguists have attempted to draw a parallel between the structure linking sentences within a text and the structure linking words within sentences. The notion of story-grammars, or text-grammars, grew out of this analogy, leading to the representation as context-free rules of the regularities appearing in such simple texts as fables.

Rumelhart (1975) describes a story as an introduction followed by episodes. An episode is an event followed by a reaction. A reaction is an internal response followed by an overt response, etc. A simple observation supporting the parallel is that two sentences in sequence usually bear some kind of relation to each other (often implicit); otherwise, the juxtaposition would be somewhat bizarre. Recognizing a paragraph as a sequence of sentences "at a syntactic level" leads to building a tree structure that may be further used by a semantic component.

The limits of the analogy between phrase structure and text structure can be easily ascertained. Winograd (1977) underlines the limits of a generative approach by pointing out that "there are interwoven themes and changes of scene which create a much more complex structure than can be handled with a simple notion of constituency." Furthermore, even if one can give an exhaustive list of words satisfying <NOUN>, it is difficult to determine how to match a <CONSEQUENCE> or an <OVERT-RESPONSE>. It follows that whether or not the process of a grammar rule has been satisfied is not easy to define. Even if we can predict that a determiner will precede an adjective or a noun, it is much more difficult to foresee that an *emotion* will be followed by a *reaction*, or at least not with the same regularity. It also seems that the "syntactic" category of a phrase is strongly domain-dependent. A given sentence may be a *consequence* or a *reason* according to the context. This phenomenon occurs less frequently with traditionally syntactic categories.

In addition, flashbacks are commonly used when people tell stories. In particular, a consequence might very well precede an explanation of an event. Chronological order is not often respected, as in "Van Gogh had difficulties to wake up. He had drunk a lot the night before." Along the same lines, elliptical phenomena (incomplete inputs) seem difficult to resolve; if one can determine the missing part of a sentence by reference to the syntactic structure of the preceding sentence, it is not easy to guess the nonstated event that has caused a reaction. The "syntactic" categories of text-grammars correspond more or less to schemas. The model defined in BAOBAB merely defines a partial ordering, or links of a preferred ordering between schemata. It follows that the *introduction* may be absent or that *signs* may precede *symptoms* without the text being regarded as incoherent. Violations of the idealistic model only cause requests for clarification or additional information. They make the comprehension process more difficult but do not halt it.

33.5 Direction for Future Work

The BAOBAB grammar is currently not very large (about 200 rules), and only seven schemata have been implemented. In order to be able to parse more complicated texts efficiently, ones involving symptoms that might imply different infections (with interactions between them), a computer program will need more sophisticated clues to determine which schema is the most appropriate to apply and in what priority. Furthermore, certain concepts are currently ignored because their relevance to medical knowledge is not always straightforward. For example, an infection acquired by a member of the patient's family or even the patient's occupation could give hints to the physician during the early stages of the diagnostic formulation, but these are not currently used.

33.6 Sample Sessions

The following examples demonstrate BAOBAB's capabilities as a natural language interface for describing cases to MYCIN. The physician types the input in upper case after the double asterisks; BAOBAB's responses are in lower case. An actual summary written by a physician has been chosen. It is slightly simplified, broken down into inputs, and given as a dialogue to facilitate the interactions. A few italicized comments have been inserted in parentheses to clarify points. Names of schemata start with the symbol \$.

In the first example, the exposition of the summary does not satisfy the model (mainly because expectations are not fulfilled), which therefore causes interventions from the program. Questions are asked, analogous to the way in which MYCIN asks them. In the other two examples, the model is satisfied; the dialogue follows a normal course.

33.6.1 First Example

** A 19 YEAR OLD BOY WITH PETECHIAL RASH, LETHARGY AND POSITIVE BRUDZINSKI'S SIGN.

Patient's name?

(The \$DESCRIBE schema has been activated by the beginning of the input. Then there is a schema-shift which generates the request for the missing name.)

** BILL

My understanding is:

The name of the patient is Bill
 The age of the patient is 19 years
 The sex of the patient is male
 Petechial is one of the types of rash which the patient has
 Lethargy is one of the symptoms that the patient had
 Brudzinski is one of the signs that the patient showed

** HE HAD AN LP

My understanding is:

A lumbar puncture has been performed on the patient

** BYE

(Unfulfilled expectations trigger requests for CSF results.)

What is the total WBC count of the CSF?

** 9500

What percent of the WBCs in the CSF were polys?

** 70

What is the protein value (mg/100cc) in the CSF?

** 32

What is the glucose value (milligrams percent) in the CSF?

** 12

33.6.2 Second Example

** REASONS FOR ADMISSION: CHILLS HEADACHE MALAISE

(sets up expectations)

So I will expect chill headache malaise as values for symptoms

** DATE OF ADMISSION: DECEMBER 25 1978

My understanding is that:

the patient was admitted at the hospital 4 days ago

** HISTORY OF THE PRESENT ILLNESS

(announces the \$IDENTIFICATION schema)

Ok.

** PAT IS A 26 YEAR OLD CAUCASIAN FEMALE

My understanding is that:

the age of the patient is 26 years

the race of the patient is Caucasian

the sex of the patient is female

the name of the patient is Pat

** ONE DAY PRIOR TO ADMISSION SHE EXPERIENCED
 CHILLS AND HEADACHE

(shift to \$SYMPTOMS)

My understanding is that:

chill is one of the symptoms that the patient had 5 days ago

headache is one of the symptoms that the patient had 5 days ago

(time inferences)

** ONE DAY AFTER SHE HAD A MALAISE

My understanding is that:

630 Strategies for Understanding Structured English

malaise is one of the symptoms that the patient had 4 days ago

** A LUMBAR PUNCTURE WAS PERFORMED

(weak clue to \$LABDATA)

My understanding is that:

a lumbar puncture has been performed on the patient

** SHE DENIED SYNCOPE OR PHOTOPHOBIA

(\$SYMPTOMS and \$SIGNS)

My understanding is that:

syncope is not one of the symptoms that the patient had

photophobia is not one of the symptoms that the patient had

** SOCIAL HISTORY: OCCASIONAL ALCOHOL USE

Do you suspect recent alcoholic history in Pat?

(clarification)

** NO

My understanding is that:

the patient is not an alcoholic

** PHYSICAL EXAM

(announces \$EXAM)

Ok.

** WEIGHT 110 POUNDS

My understanding is that:

the weight of the patient is 49.9 kgms

** TEMP 101 DEGREES P 90 BP 110/75 RR 28

My understanding is that:

the temperature of the patient is 101

the pulse of the patient is 90

the blood pressure of the patient is 110/75

the respiratory rate of the patient is 28 breaths/minute

** SKIN NORMAL

(leads to a clause via inference)

My understanding is that:

the patient does not have a rash or cutaneous lesions

** NECK: MARKED NUCHAL RIGIDITY

(another inference)

My understanding is that:

meningismus is one of the signs that the patient showed

** BRUDZINSKI'S SIGN WAS POSITIVE

My understanding is that:

Brudzinski is one of the signs that the patient showed

** LAB DATA

(announces \$LABDATA)

Ok.

** CSF RESULTS

(sets a context)

Ok.

** WBC 9500

(understood in context)

My understanding is that:

the WBC from the CSF is 9500

** PROTEIN 56

My understanding is that:

the CSF protein is 56

** GLUCOSE 45 MG%

My understanding is that:

the CSF glucose value is 45

** PERIPHERAL CBC *(sets another context)*

Ok.

** WBC 23.5 *(understood in context)*

My understanding is that:
the white count from the patient's peripheral
CBC (in thousands) is 23.5

** THE VALUE OF WBC IS 22.5 *(corrects a mistyping)*

Instead of 23.5 ? *(making sure that it was meant)*

** YES

My new understanding is that: The white count from
the patient's peripheral CBC (in thousands) is 22.5

** BYE

I have assumed the following by default: *(inferences by default)*
the state of consciousness of the patient is alert
the patient's X-ray is not abnormal
the patient is not immunosuppressed
Do you agree?

** THE PATIENT IS IMMUNOSUPPRESSED *(overrides a default value)*

My new understanding is that: The patient is immunosuppressed
Do you agree?

** YES

bye.

33.6.3 Third Example

** REASONS FOR ADMISSION: HEADACHE, LETHARGY AND FEVER

So I will expect HEADACHE LETHARGY FEBRILE as symptoms

** DATE OF ADMISSION ON MARCH 14

That was 6 days ago.

My understanding is
the patient was admitted at the hospital 6 days ago

** BARBARA IS A 3 YAER OLD WHITE FEMALE

= YEAR *(automatic spelling correction)*

My understanding is
the age of the patient is 3 years
the race of the patient is Caucasian
the sex of the patient is female
the name of the patient is Barbara

** ONE DAY PRIOR TO ADMISSION, SHE BECAME LETHARGIC

My understanding is
lethargy is one of the symptoms that the patient had 7 days ago

** SHE COMPLAINED OF HEADACHE AND HAD A TEMPERATURE OF 101 DEGREES

My understanding is
headache is one of the symptoms that the patient had
the temperature of the patient is 101
the patient is febrile

** A PHYSICIAN OBTAINED A COMPLETE BLOOD COUNT

My understanding is
the results of a recent CBC of the patient are available

** A LUMBAR PUNCTURE WAS PERFORMED

My understanding is
a lumbar puncture has been performed on the patient

** WEIGHT 39 LBS

My understanding is
the weight of the patient is 17.7 kgms

** HEIGHT 1.03 METER

Is the patient really 103.0 centimeters tall? *(asking for confirmation because that seems tall for 3 years)*
(Y or N) ** YES

My understanding is
the height of the patient is 103 cms

** BP 120/70 P 140

My understanding is
the blood pressure of the patient is 120/70
the pulse of the patient is 140

** RR 28 T 101

(redundancy, the temperature was already known)

My understanding is
the respiratory rate of the patient is 28

** SKIN: PETECHIAE ON FACE AND SHOULDERS

My understanding is
petechial is one of the types of rash which the patient has

** NECK: DEFINITE MENINGISMUS

My understanding is
meningismus is one of the signs that the patient showed

** LAB DATA: CSF RESULTS

(announces lab results; no clause is built)

Ok.

** WBC 110 WITH 96% POLYS

(interpretation in the context of CSF)

My understanding is
the WBC from the CSF is 110
the percent of PMN's in the CSF is 96

** PROTEIN 25 MG% AND GLUCOSE 95

My understanding is
the CSF protein is 25
the CSF glucose value is 95

** PERIPHERAL CBC

Ok.

** WBC 12300

Is the white count from the patient's peripheral CBC (in thousands) really 12300?
(the physician forgot that a number of thousands was expected)

** NO

Please enter the right value now

** 12.3

My understanding is

the white count from the patient's peripheral CBC (in thousands) is 12.3

** WITH 76% POLYS AND 1% BANDS

My understanding is

the percent of PMN's in the CBC is 76

the percent of peripheral WBC's which are immature in the CBC is 1

** SIMULTANEOUS BLOOD GLUCOSE: 170

My understanding is

the simultaneous blood glucose value is 170

** BYE

I have assumed the following by default:

the state of consciousness of the patient is alert

the patient has not shown symptoms of mumps

it is unknown whether the patient's X-ray is abnormal

the patient does not have a rash or cutaneous lesions

the patient does not have evidence of ocular nerve dysfunction

the patient is not an alcoholic

the patient is not immunosuppressed

it is unknown whether the patient does have a TB risk factor

the patient is not a compromised host

the patient does not live in a crowded environment

the patient has not been recently exposed to a contagious disease

the patient has not undergone surgery

the infection was not acquired while the patient was hospitalized

the patient has not been seriously burned

there is not evidence of congenital malformation involving the central nervous system

the patient is not a diabetic

the patient is not a premature infant

it is unknown whether the patient has had an injury or insult to, or defect in the central nervous system

You can change any of them if you disagree

** SHE HAS UNDERGONE SURGERY

I suppose it should replace the patient has not undergone surgery

Right?

** YES

Ok.

** GOOD-BYE

33.7 Conclusion

The strategies outlined in this chapter could be applied to a broad range of structured texts. The approach rests on the assumption that the texts' scenarios can be seen as sequences of episodes, identifiable by the program, in order to be integrated into appropriate schemata. Therefore, clustering attributes into framelike structures must make sense in the domain of application. The episodes could simultaneously refer to several schemata; that is, the associated schemata could have slots in common. Furthermore, it should be possible to define partial-ordering links between schemata.

The relationships could be rather loose, but the more constrained they are, the better this feature would work.

Expert systems usually need some kind of understanding to communicate in natural jargon with their users (expert, consultant, and/or student). The technique described here—breaking the knowledge down into schemata that correspond to different pieces of texts, associating semantic grammar rules with the schemata, and using strategies for recognizing episode shifts—should be generally applicable in such domains.