SRI International

A TEAM USER'S GUIDE

Technical Note 254

21 December 1981

By: Armar A. Archbold, International Fellow
Barbara Grosz, Senior Computer Scientist
Daniel Sagalowicz, Assistant Director

Artificial Intelligence Center
Computer Science and Technology Division

Contract N00039-80-C-0645
DARPA Order No. 3988
SRI Project 1605

Prepared for:

Defense Advanced Research Projects Agency
Information Processing Techniques Office
1400 Wilson Boulevard
Arlington, Virginia 22209

Attention: CDR R. Ohlander

SRI International®

333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 859-6200 • TWX: 910-373-2046 • Telex: 334 486

Preface

TEAM (Transportable English Data Access Manager) is a computer system designed to acquire information about a (local or remote) database, and subsequently to interpret and answer questions addressed to the database in a subset of natural language.

The system currently includes an interactive program that acquires information about a database from the database administrator. The user is asked how the database is to be structured and what language is to be used to talk about its contents. The program augments the following components of a core natural language access system:

* The lexicon
* The conceptual model of the domain (or conceptual schema)
* The structural model of the database (or database schema).

After these steps have been completed, the system accepts natural-language questions about the database contents and, to the extent it is able, provides relevant responses.

Current Limitation: At present the TEAM system allows interrogation only of local databases that have been created by the user. When fully developed, it will also allow the description and interrogation of remote or "external" databases, whose content the user has not furnished

and cannot alter.

Throughout this manual you will find important warnings about current limitations of the system, as it is undergoing constant development and is therefore essentially incomplete.

These warnings are prefaced by "**Current Limitation:**".

The number of English expressions and constructions that can be described to the system and then utilized to interrogate it is being steadily expanded. Examples given in this manual are thus suggested, not mandatory forms. Users are encouraged to experiment with different constructions, because that is how omissions in TEAM's coverage can be exposed (and, preferably, remedied).

If you have any questions or comments, please direct them to SAGALOWICZ at SRI-AI or GROSZ at SRI-AI.

CONTENTS

I INTERACTING WITH TEAM 1

A. Starting Up TEAM 1

B. Typing Input to the System 1

C. What TEAM Can Do--Commands and Questions 2
1. Commands 2
2. Questions 3

II ACQUIRING NEW INFORMATION ABOUT THE DATABASE 5

A. Name of the New File 8

B. Names of the Fields 8

C. The File Subject 10
1. Questions About Primary Keys and "Identifying" Fields 11
2. Questions About Expressions that Refer to a File Subject 13
3. Question About Fields with Subject Name Values . 18

D. Information about Fields of the File 19
1. General "Unknown/Not Applicable" Convention for the Fields 21
2. Questions Asked About All Fields, Irrespective of Type 22
3. Questions Asked About Symbolic Fields 25
4. Questions Asked About Numeric Fields 27
5. Questions About Feature Fields 33

E. Acquiring New Verbs 39
1. File for Verbs 39
2. Names of Verbs 39
3. Past Tense of the Verb 40
4. Past Participle of the Verb 41
5. What Kind of Verb 41
6. SUBJECT OF THE VERB 42
7. Object of the Verb 42
8. The Verb Pattern 44
9. Correspondence Between Verb and Noun 44

F. End of Acquisition 44
G. Backup Mechanism During Acquisition 45

III OTHER MISCELLANEOUS FUNCTIONS IN TEAM 51
A. Database Entry 51
B. User Profile 56
C. Saving and Loading the Schemas 57
D. Miscellaneous Commands 57

IV QUERYING DATABASES 59

I INTERACTING WITH TEAM

A. Starting Up TEAM

To run TEAM, simply type
<team>TEAM.EXE

at the top level, followed by carriage return (<CR>). Once the system has started up, the user is prompted by a number followed by a left arrow (or an underscore on some terminals), e.g.:
1_

B. Typing Input to the System

When TEAM prompts the user, he should type in his question or command, terminating it with a carriage return. In general, TEAM will interpret any carriage return as the end of a question. If you are typing a question longer than one line, either do not type a carriage return until the end, or precede the carriage return with a space. Punctuation is permitted but not required. TEAM accepts input in any combination of upper- and lowercase. Backspacing is done by typing CTRL/A or the DELETE key, depending on whether the host operating system is TENEX or TOPS-20, respectively. You may abort the processing of a question by typing CTRL/D. Further procedures for editing a line of input are described in Section 14 of the INTERLISP manual [1].

Current Limitations--The current implementation is not as robust as we would like: Certain punctuation may cause errors; if this happens, try the same question without any punctuation. Also, typing CTRL/D may damage the system irreparably because of problems caused by its implementation in three separate files (virtual machine images).

C. What TEAM Can Do--Commands and Questions

The user may type to the system either (1) a key command (using predefined keywords; or (2) a question in English. Each of these capabilities is described briefly below and in more detail in Chapters II, III, or IV.

1. Commands

There are currently nine key commands that invoke procedures:

ACQUIRE	to specify structural and linguistic information about database files
VERBS	to specify information about verbs used in interrogating the database files
EDIT	to specify the contents of local files
LOCATION	to specify the locations of remote files on external database management systems
SAVE	to save the internal data structures of TEAM
LOAD	to retrieve the TEAM data structures saved in a previous session
PROFILE	to load, or specify and save, a user profile, telling the system how to respond to queries
QUIT	to end a session and exit from TEAM
HELP	to have the list of possible commands displayed

Current Limitation--Since the system currently learns about new field values only when an EDIT command is used, the user must use that

command even if a remote database exists; the user then creates a local database that includes all the field values he wants the system to interpret in queries.

The first time any section of a database is described, the normal sequence of commands is ACQUIRE, EDIT, VERBS, LOCATIONS (optional), and SAVE. Thereafter the user should start with a LOAD command. The SAVE is very important for further use of the acquisition.

The ACQUIRE and VERBS commands are described at length in Chapter 2 of this manual. The remaining commands are discussed in Chapter 3.

2. Questions

After database files have been described to the system (using ACQUIRE), and their contents made known to the system (using EDIT), the user, in answer to the prompt, can type English questions about the information contained in those databases.

For example, suppose the user has specified a local database called *SHIP*, which is concerned with the attributes of various ships and with the following structure and content:

ID	NAM	COMMANDER	SPEED	ASW
G35-T9	WHALE	J.JONES	25	Y
B92-R4	KENNEDY	T.SMITH	18	N
W12-U89	FOX	W.BRADLEY	30	Y
K65-R67	INTREDPID	P.CHARAT	22	N

The user can then ask (and receive answers to) such English questions as

2_ WHO IS THE COMMANDER OF THE FOX?
(Answer: W.BRADLEY)

3_ WHAT SHIPS HAVE A SPEED GREATER THAN 20 KNOTS?
(Answer: WHALE,FOX,INTREPID)

4_ WHAT SHIPS ARE SUB-KILLERS?
(Answer: WHALE,FOX)

For a question to be answerable, it must

* Ask for information contained in the database. For example, the system could not answer a question like "Is the Fox a U.S. ship?" because the nationalities of the ships are not specified in the above database.
* The user can use only content words about the database (essentially nouns, verbs, and adjectives) which he has explained by means of the ACQUIRE and VERBS commands. For example, "What ships are sub-killers?" will be answered only if the user has told the system that "a ship is a 'sub-killer'" means that that ship has a 'Y' value in its ASW-CAPABLE feature field.

II ACQUIRING NEW INFORMATION ABOUT THE DATABASE

To provide access in English to a particular database, the TEAM system must be given three kinds of information. Specifically, it needs to know about

(1) The English words and expressions that will be used in queries.
(2) The structure of the database it is to access--e.g., the files are in it, the kinds of objects the files describe (their subjects), the names and types of their fields, and so on.
(3) The content of the database--that is, the contents (values in the fields) of local or "internal" database files. (See Chapter 3.)

Information of types (1) and (2) can be provided to the system by means of the "acquisition" procedures, which are initiated by typing the ACQUIRE and VERBS commands to the system's prompt. These acquisition procedures are the subject of this chapter. Information of type (3) is provided through the editor (invoked by the EDIT command) as described in Chapter III.

Both the ACQUIRE and the VERBS commands initiate a series of questions about the structure of, and the language used to interrogate a particular <u>single</u> file. Sections 2.1 through 2.5 of this chapter describe the questions asked after typing ACQUIRE, while Section 2.6 describes those that result from typing VERBS. Section 2.7 describes

the final questions in each of these series.

The user should note the following features of interaction with the acquisition procedures:

- Acquisition dialogues concern one and only one file in the database at any one time. In other words, after typing ACQUIRE or VERBS to the top-level prompt, one can only specify information (or modify information previously specified) describing a single file. To specify information about several files, one must type the ACQUIRE or VERBS commands as many times as there are files one wishes to work on.
- Each question asked by the system is initially presented in a rather concise format, which we refer to below as the question's "short form."
- By typing a question mark (?<cr>) in response to the short form of any question, the user will be presented with an expanded, "long form" of that question, which provides more of an explanation of what the system requires in the way of a specification. Changes are bound to occur while the system is undergoing further development. The user is advised to make frequent use of this ? explanation mechanism to get help.
- The user can type two different kinds of answers to some of the acquisition questions: "words" and/or what is referred to in this manual as "multiwords."

  A word is an unbroken string of alphanumeric characters, containing no spaces (i.e., blanks), and whose first character is not a number. Thus SHIP, *SHIP*, ID and G34-T6 are words, whereas 345, ID NUMBER, and 3G34-T6 are not. A multiword, on the other hand, is a parenthesized list of words separated by spaces. Thus (ID NUMBER) is a multiword.

  The user will normally type in words, like SHIP. However, one might wish, in response to certain questions, to type in both words and phrases, i.e., standard sequences of words separated by blanks. Such sequences of words separated by blanks must be typed in as a parenthesized list--as a multiword--so that TEAM can recognize that those words "belong together." For example, if the user wants to specify BOAT, PLATFORM, and NAVAL VESSEL as synonyms for SHIP, the phrase (multiword) NAVAL VESSEL must be parenthesized. We would have the following:

```
Synonyms for SHIP? >  BOAT PLATFORM (NAVAL VESSEL) <cr>
--------------------  --------------------------
  !                     !
  Question asked by TEAM   User's answer: 2 words,
                                          1 multiword
```

Multiwords can only be synonyms; the subject of the file, field names, and field values must be words.

- The program expects the user's responses to its questions to have a certain form and content. Should the user's answers not meet the system's expectations, the program will advise the user as to what, in its view, is unsatisfactory about the response, and will ask the question once again. In some cases, the system will display its expectation as to what a "normal" or "default" answer to a question should be. The user can then either accept the default (by typing Y<cr>), or refuse the default (by typing N<cr>) and then typing in the desired answer.
- The system asks acquisition questions in a structured sequence. However, a backup mechanism provides the user with a means of correcting previous responses, and hence with control over the flow of the acquisition dialogue. Almost any time during acquisition, the user may type !<cr> (in response to a question) to change his answer retroactively to any previous query.

  Examples of the backup mechanism will be given in section 2.7.

  After typing !, the user can either (1) type a keyword that appeared in a previous TEAM question--or in a previous user's answer to a TEAM question--that the current user wants to reconsider; or (2) type !! and be shown all the questions presented so far with their previous answers--with the opportunity to state which changes he wishes to make.

The next sections give examples of each of the kinds of acquisition questions posed by TEAM. The questions will be shown in both their "short" and "long" forms. We shall also occasionally describe the general nature of the information requested when this is not obvious from the question itself.

A. Name of the New File

Short form: File name ?

Long form: The name of the new file is (type file name) ?

This is the first question asked when the user has typed ACQUIRE and has opted to provide information about a new file.

The answer should be a single word, like *SHIP*, not a multiword like (NAVAL SHIP). The word is the proper name of the file, a proper name that is unique to that file. <<No two files can have the same name>.

The system has implicit knowledge that the answer is the name of a file in the database.

For the rest of the examples below, we shall assume that the user's answer is *SHIP*. In other words, we assume the following:

```
File name ? >  *SHIP* <cr>
-------------- ----------
 !                 !
 question asked    user's response
 by the system
```

B. Names of the Fields

Short form: *SHIP*'s fields ?

Long form: The field names in file *SHIP*
are (type sequence of column headers) ?

This question asks for the headers of the columns in the previously declared file (in our sample dialogue, *SHIP*).

The user's answer should be the names of the columns in the the file. Each of these names should be a word, like UICVCN, COMMANDER, HOME-PORT, LENGTH. Multiwords like (HOME PORT) are prohibited. These column names must be in the singular form (e.g., HOME-PORT rather than HOME-PORTS). Abbreviations (e.g., "LGH" instead of length) are, of course, acceptable.

Unlike the previous case of the file name, the system knows what these are only in the most general terms (lexical items and column headers); further information about their syntactic and semantic/conceptual nature must be elicited by subsequent questioning.

For the purposes of the examples below, we shall assume that the user is specifying information about a file *SHIP*, which records the name (abbreviated to the column header NAM), the identification-number (abbreviated ID), the commander (COMMANDER), the speed (SPEED), and the antisubmarine warfare capability (ASW) of each of a number of ships. That is, the user is supposed to have in mind a file *SHIP* of the following configuration:

File: *SHIP*

ID	NAM	COMMANDER	SPEED	ASW
G35-T9	WHALE	J.JONES	25	Y
...	...	...	...	...

We thus assume that the response to this question is as follows:

```
*SHIP*'s fields ? > ID NAM COMMANDER SPEED ASW <CR>
-------------------  ------------------------

  !                          !
  question asked by          user's response
  the system
```

C. The File Subject

In this section, we describe a series of questions the system asks about the primary keys of the file. This block of questions allows the user to inform the system about the following:

* Which fields of a file are its primary keys.

  Primary keys are those fields whose values uniquely identify one of the objects the file is about; thus, in our sample *SHIP* file concerning ships, the value of the ID field in each row identifies a ship whose attributes are specified by the values in the other fields (columns) in that row. The values in primary keys are unique identifiers because, by definition, no two values are the same.

* Which fields are "conveniently identifying" keys to be returned to the user for the purpose of identifying entities when answering questions.

  In some files there are fields other than the primary keys whose values are convenient, though not unambiguous, identifiers of the objects described by the file. Thus, in our sample *SHIP* file, although the values in the primary key ID field are unambiguous names for each ship described by the file, the values in the NAM (ship name) field are helpful ways of identifying the ships to the user in answering questions, even though two ships with necessarily different identification numbers might share the same name.

* Which words (nouns like "ship" and "vessel," and pronouns like "it" and "she") can refer to the subject of the file.

* Which field contains values that are names of the file subject.

1. Questions About Primary Keys and "Identifying" Fields

a. Primary Keys for a Subject/File

Short form: Primary keys(s) of *SHIP* ?

Long form: Which of the following fields of *SHIP*
(NAM ID COMMANDER SPEED ASW)
are the file SHIP's primary keys ?

That is, which are those fields whose values individuate
the entities the file is intended to describe?

Type the names of one or more primary key fields.

The phrase "the entities the file is intended to describe" will be used again later to ask for the subject of the file.

Current Limitations and Assumptions--At present the TEAM system assumes that

(1) Each file has only one primary key.

(2) The primary key of a file is always a symbolic field, rather than an arithmetic or feature field.

Roughly, this means that the values in a primary field are words, not numbers (for a more extended discussion of types of fields, see Section 2.4 below). So a primary field, like the ID field in our *SHIP* sample file, could have as values only alphanumeric strings starting with a nonnumeric character, such as G35-P5 or #38450, but not a number like 38450.

We assume that the answer is ID in our sample *SHIP* file:

File: *SHIP*

ID	NAM	COMMANDER	SPEED	ASW
G35-T9	WHALE	J.JONES	25	Y
...	...	...	...	...

The ID field is the field whose values unambiguously individuate the objects described by the file *SHIP*.

b. Convenient 'Identifying' Keys for a Subject

Short form: Convenient "identifying" field(s) ?

Long form: Are there fields of *SHIP* among those declared
(NAM ID COMMANDER SPEED ASW)
whose values you would like the system to use, when responding to your questions, to identify entities described by the file (even if those values might not strictly individuate entities)?

Please type a sequence of field names or <cr>.

This question determines whether there are some fields whose values the user would like to have returned in responses that require naming individual file subjects. For example, a user might refer to employees by name rather than by ID number, even though two employees with inevitably different IDs might share the same name. The identifying field values do not need to be unique. ID are used to make answers more ***** to the user.

For purposes of the example, we assume that the user answers NAM here. In our sample *SHIP* file:

File: *SHIP*

ID	NAM	COMMANDER	SPEED	ASW
G35-T9	WHALE	J.JONES	25	Y
...	...	...	...	...

the user would then want to have the ships identified by values in the NAM field, rather than by the ships' ID numbers.

2. Questions About Expressions that Refer to a File Subject

a. Questions About the Name of the File Subject

Short form: Name of this file *SHIP*'s subject (use singular) ?

Long form: Please specify a word for the subject of the file *SHIP*, i.e., for the entities individuated by the values of the primary key fields in the file and described by the values of the nonprimary fields in the file.

A word wwww for the subject of the file is such that one can properly say that

the file *SHIP*

gives information about each wwww
in a set of wwww's

For example, the file PERSONNEL gives information about each EMPLOYEE in a set of EMPLOYEES.

(Please note that the singular form of the word wwww is required--i.e., EMPLOYEE rather than EMPLOYEES.)

Please specify one word for the subject of the file. >

This question asks for the noun that designates the class of the entities individuated by the primary key fields and described by the nonprimary fields of the file. It is important because the reply to it creates the lexical and conceptual information to which all the nonprimary fields will be linked.

The answer must be a single word, not a multiword.

Furthermore, the system wants a singular "count noun"; care must be exercised by the user to provide an appropriate one. Roughly defined, a count noun is a noun that can appear with both the definite article

"the" (e.g., "the ship") and the indefinite article "a/an" (e.g., "a ship") and can appear in either singular or plural form, preceded by a number term (e.g., "the ship," "the ships," "one ship," "two ships"). Occasionally some effort must be made to provide the system with such a noun for the subject of a file. To illustrate the potential difficulty, imagine that we have a file about cows and bulls:

NAM	ID	PRICE	AGE	SEX
TORO	#34-T5	3000	2	M
BETSY	#49-R4	2250	1	F
...	....	....	...	...

Though it might be natural to infer that the subject name of such a file is CATTLE or LIVESTOCK, these are not acceptable nouns for the TEAM system: one cannot speak grammatically of "a cattle," "a livestock," "one livestock," "two livestock," and so on. One would therefore have to find or devise some word that qualifies as a singular count noun, such as "bovine-animal."

b. Unconventional Plurals of Subject Names

We assume that the response to the preceding question was SHIP--that is, that the subject name of the illustrative *SHIP* file is SHIP. Then the next question will be

```
Short form: Plural of SHIP?
 = [Default: SHIPS ]    Default accepted (Y,N)?
```

This question has no long form. If the default SHIPS is not

accepted by typing "y" as an answer, the irregular plural must be supplied. Thus, in the example the user would simply type Y and accept the default plural SHIPS. However, if the subject name were KNIFE, the system would offer the incorrect default plural KNIFES and the user would type N, thereby rejecting the default. He would then type the correct form KNIVES.

c. Questions About References to a Subject

TEAM must know which pronouns can refer to file subjects. Therefore TEAM currently asks the user about (1) the "apparent gender" of the file subject, and (2) its "humanity." "Apparent gender" determines whether one can refer to something as "he," "she," or "it" (him/himself/his, her/herself/hers, itself/its, etc.). It is called apparent because the correspondence between the biological category "sex" and the grammatical category "gender," albeit very close, is far from absolute:

Noun	Referential Pronoun	Examples
a doctor	he, she	
a mother	she	
a baby	he, she, it	
a tiger	it (sometimes: he, she)	
a fish	it (rarely: he, she)	
a ship	it, she	How fast does she sail?
France	it, she	It is one of the largest countries in Europe. France has been able to increase her exports by 10%.

"Humanity" determines whether the pronouns "who" or "which"/"what"

can refer to the subject. By and large, "who" applies to human beings (or perhaps to the higher animals), "what"/"which" to the lower animals and inanimate objects. However, there are some exceptions.

Noun	Referential Pronoun
a mother	who
a baby	who, which/what
a family	who, which
a tiger	which/what
a fish	which/what
a ship	which/what
France	who, which/what (country)

Current Limitation--There is a third aspect of referring to file subjects that is not now being determined by the TEAM system, i.e., the "number" of the noun for the subject. The number of a noun usually governs the singularity or plurality of its attendant verb forms. Normally, if a count noun has both a singular and a plural form, the former will take a singular verb and the latter a plural verb. But collective nouns can take pronoun substitutions that are either singular (it) or plural (they) without any change of number in the noun (the family: it/they; the families: they). This is true of some collective nouns with unique reference that are very much like proper nouns:

(1) A family was sent to Indonesia.
They soon found a home in the capital.

(2) Congress was in session.
They were deliberating furiously.

Because TEAM does not currently take account of the number of nouns, it will mistakenly regard certain grammatical sentences as being ungrammatical. Thus, it would not recognize "The family are downstairs" as a well-constructed sentence.

d. Pronouns for Subjects

Short form: Can you refer to a SHIP using (1) he, (2) she, (3) it, or (4) they?

Please type in all the pronouns you could use.

Long form: Which of the following pronouns could be used to refer to a SHIP ?

(1) he
(2) she
(3) it
(4) they

Please type in all the pronouns you could use.

Assume that the user's answer here is "IT SHE", since often "A SHIP is in the Mediterranean." One can say "IT is a destroyer," or "SHE is a destroyer."

e. Humanity of a Subject

Short form: Is SHIP human (Y or N) ?

Long form: Is a SHIP human? Namely, can one refer to a subject of this file *SHIP* by using who rather than what?

Please answer Y or N.

Normally a ship is not regarded as a human entity: we would say "The ship that is in the Mediterranean" rather than "The ship who is in the Mediterranean," so the most plausible answer would be N. However,

sometimes users might ask questions like "Who sends radio messages?", meaning "Which ships send radio messages?", or, more explicitly, "People on board which ships send radio messages?". Such users should answer the above query with Y.

f. Synonyms for the Subject Name

Short form: Synonym(s) for subject name SHIP?

Long form: You said that you will refer to an individual subject of this file SHIP using the word SHIP.

Are there other similar words (i.e., synonyms of
SHIP)
that you would also like to use to refer to a subject of this
file?

If not, type <cr>; if so, please type a list of those similar words.

The user then inputs this list of synonyms (for example: BOAT PLATFORM (NAVAL VESSEL) <CR>). For each such synonym, TEAM would then ask the question previously asked about the irregular plurals. For example:

Plural of (NAVAL VESSEL)?
= [Default: (NAVAL VESSELS)] Default accepted (Y,N)? Y <cr>

3. Question About Fields with Subject Name Values

Short form: Field specifying the name of each SHIP ?

Long form: Each value in the primary key field(s),
(ID)
individuates some SHIP.

Is there another field whose corresponding value is the name
of that same SHIP?

If not, type <cr>; if so, please type that single field name.

(In our example, the user answers NAM at this point.)

This concludes the category of questions related to the primary key fields and the subject of the file.

D. Information about Fields of the File

General Information Required

The fields of a file typically specify either properties of the entities individuated by the primary keys of the file or other objects with which those entities enter into certain relations. Fields are broken down into three distinct types:

(1) Symbolic fields, whose values are usually nouns (proper nouns, including names in the intuitive sense and "numeric strings" that are "rigid designators," such as employee ID numbers, count nouns, and mass nouns) or adjectives; these values do not permit such real arithmetic operations as division and averaging, although they may be associated with some scale and may allow ordinal calculations (e.g., the values in a MILITARY-RANK field).

(2) Arithmetic fields, whose values--"pure numbers" or numeric measures expressed in some unit of measure--may not only be added, subtracted, etc., but also multiplied, divided and averaged; dates, which are a special case, also fall into this type.

(3) Feature fields, whose values are either Boolean (two-valued, e.g., T/F), showing that some attribute is true or false of some object described by the file, or else n-valued (e.g., CERTAIN, PROBABLE, IMPROBABLE...) showing degrees of likelihood that some object has some particular attribute.

Examples of the kinds of values expected in these types of fields are as follows:

SYMBOLIC
========

Headers: nouns of various types; for example:
NAME, EMPLOYEE-ID-NUMBER, CLASS,
COUNTRY, NATIONALITY,
COMMANDING-OFFICER, TYPE, WEAPONS,
CARGO

Values: nouns of various types, and adjectives
19861-90, (proper nouns, numeric
109-A7-64P strings, "rigid desig-
nators," but not names)
J.F.KENNEDY (proper noun, name)
SUBMARINE (count noun)
WHEAT (mass noun)
FEET (unit/measure noun)
BRITISH (adjective)

ARITHMETIC
==========

Headers: nouns (sometimes with associated
nouns for units of measure,
adjectives and their comparative
and superlative forms)

Values: true numbers, measures, and dates
32 ("pure number"--e.g., as a value of a
NUMBER-OF-PEOPLE-IN-DEPARTMENT field;
cardinality of set of objects of a
certain type, but not associated with
a unit: "<number> OF people," but
*"<number> IN people," *"<number>
of people in <unit>s")
117 ("measure," e.g., as a value of a
DRAFT (of ship) field, associated
with a measure unit: "draft (of a
ship) in feet")
9/12/62 (date--as value of LAUNCH-DATE)

FEATURE
=======

Headers: nouns and/or adjectives
SPOTS (noun--in a file about plants)
LEAFY, NUCLEAR-POWERED (adjectives)

Values: binary standard entries, e.g.
YES/NO, +/- (binary, for ASW-CAPABLE field)
or n-case standard entries, e.g.
CERTAIN / PROBABLE / POSSIBLE / UNLIKELY /
etc. (for MALIGNANCY field in a
file about cancer patients)

Current Limitations--At present the TEAM system cannot handle

- Symbolic fields whose values are ranked by the use of an associated scale (e.g., MILITARY-RANK);
- Feature fields with n-valued entries indicating degrees of likelihood.

1. General "Unknown/Not Applicable" Convention for the Fields

The system proposes a default set of unknown/not applicable notational conventions for fields of the various types. If the user does not accept these conventions, he will be asked if he wishes to specify different conventions for all fields of each type. Should the user not want to adopt uniform conventions for all fields of a given type, he will subsequently be asked to specify conventions individually for each field of that type.

Short form: Accept our default "unknown"/"not applicable" conventions? (Y or N)

Long form: Our default notational conventions for "unknown" and for "not applicable" values in ALL fields of each of the three possible types (arithmetic, feature, and symbolic) are as follows:

For each field of type --! V			
Notational Convention for	ARITHMETIC	FEATURE	SYMBOLIC
UNKNOWN	-1	*	*
NOT APPLICABLE	-2	**	**

Do you accept the above default notational conventions for the fields in this file *SHIP* (Y or N) ?

Let us suppose that the answer here is Y. If the answer were N, however, the following dialogue would be initiated.

Short Form: Delay specifying NK/NA notations for fields of type

```
ARITHMETIC until asked about each field? (Y or N)
>
```

Long Form: It is assumed that you will use the same conventional notation for "unknown" and "not applicable" values in all fields of a similar type. You may, however, delay your decision for the entire field of a given type, and specify the conventional notation only when you are asked about each of these fields.

```
  Please specify these conventional notations below.

What notation will you use in ARITHMETIC
fields for NK, NA?

Delay the decision (Y or N) ?
> Y
==================================================
```

Short Form: Delay specifying NK/NA notations for fields of type FEATURE until asked about each field (Y or N) ?

```
> Y
==================================================
```

Short Form: Delay specifying NK/NA notations for fields of type SYMBOLIC until asked about each field (Y or N) ?

```
> N
==================================================
```

Short Form: Uniform conventional notation for UNKNOWN in all SYMBOLIC fields of this file ?

```
> NK
==================================================
```

Short Form: Uniform conventional notation for NOT APPLICABLE in all SYMBOLIC fields of this file ?

```
> NA
```

2. Questions Asked About All Fields, Irrespective of Type

a. Type of the Field

Short form: Is NAM arithmetic, feature, or symbolic (A/F/S) ?

Long form: Is NAM

A an arithmetic field, whose values are either dates or numeric values with which it makes sense to perform operations of real arithmetic (e.g., multiplication and averaging), or

F a feature field, whose values are either Boolean (e.g., YES/NO, T/F, POSITIVE/NEGATIVE), or words indicating degrees of likelihood (e.g., CERTAIN , LIKELY, IMPROBABLE, etc.), or

S a symbolic field, whose values are strings of letters and/or numbers (e.g., words, the names or ID numbers of objects or people), that may possibly be compared in some way (e.g., the specifications of military ranks), but which do not allow meaningful arithmetic calculations like averaging and division)?

Please type A, F, or S

b. Plural Form of the Field Name

The system asks for the plural form of each field name (except those of feature fields). The question is identical to the one above about the plural form of the subject name. For example:

```
Plural of NAM?
  = [Default: NAMS ]  Default accepted (Y,N) ?
```

c. Synonyms for the Field Name

TEAM requests synonyms for each field name (except those of feature fields). The question is identical to the one above about the synonyms of a subject name. For example:

```
Synonym(s) for NAM?
>
```

to which the user might respond with "NAME <cr>".

Note that all words which the user wants to allow to refer to the

information in this field should be included here. The response might therefore have been "NAME SHIPNAME (SHIP NAME)."

d. Plurals of Field Name Synonyms

The system will ask for the plural form of each field name synonym specified. For example:

```
Plural of NAME?
  = [Default: NAMES ]  Default accepted (Y,N) ?
```

e. Conventions for Unknown/Not Applicable Values in the Field

In those cases in which the user has not accepted the general default conventions proposed by the system and, moreover, has opted to "delay" his decision about the "unknown/not applicable" values until queried about each field of a given type, the system will ask questions like the following:

```
Short Form: UNKNOWN
conventional notation for this field NAM ?

Long Form: Please type the database values of this field
NAM which stand for UNKNOWN.

  notational convention for UNKNOWN?

==================================================

Short Form: NOT APPLICABLE conventional notation for this field
NAM ?
```

3. Questions Asked About Symbolic Fields

a. Are Field Values Names of Units?

Short form: Are field values of NAM units of measure (Y or N) ?

Long form: Are the values in this field, NAM,
the names of units of measure of some kind?

For example, a CARGO-UNIT field, describing the units in which quantities of cargo (specified in a second CARGO-QUANTITY field) were expressed, would contain names of units of measure like POUNDS, CUBIC-FEET, METRIC-TONS, BUSHELS, etc.

Please answer Y or N.

The question is asked at this juncture because there may well be files with (for example) the following kinds of fields:

SHIPS

SHIP-NAME	CARGO	CARGO-QUANTITY	CARGO-UNITS
Fox	Oil	23,456	barrels
Adams	Wheat	987,716	bushels
Whale	Iron Ore	56,897	metric tons

One of these fields may contain the names of the units of measure in which another numeric field's values are expressed.

Current Limitation--TEAM does not now attempt to handle the connections between such fields as CARGO, CARGO-QUANTITY, and CARGO-UNITS in the foregoing example. TEAM will eventually have to handle such intrafile linkages. At present only syntactic information is extracted from the user's response to this question.

b. Are Field Values Proper Names of Field Domain Objects?

Short form: Are field values proper names of each NAM (Y or N) ?

Long form: Is every value in this field NAM
a proper name of some individual NAM ? For example:

> Each of the values in a COUNTRY-OF-ORIGIN field, values such as ENGLAND, FRANCE, U.S.A., etc., would be proper names of some COUNTRY-OF-ORIGIN.
>
> The values in a CARGO field, values such as WHEAT, OIL, PASSENGERS, etc. are NOT proper names of CARGOs.
>
> Note: the values in a SHIP-NAME field, values such as WHALE, FOX, BISMARCK, etc., are also NOT names of particular SHIP-NAMEs (i.e., names are names of objects, not names per se).

Please answer Y or N.

c. Uniform Humanity of Field Values

Short form: Are NAMs human?

Long form: Are NAMs human?
Can one ask "Who is the NAM? ?"

Please answer Y or N.

d. Field Values as Modifiers of the Subject

This question is posed in two stages:

(1)

> Short form: Typical value in NAM ?
>
> Long form: Please type in a typical example of the values that would appear in this field
>
> Let us suppose that the user types in FOX as an example of the values in the NAM field. Then the following question would be asked:

(2)

Short form: Can NAM values be used as modifiers (Y or N)?

Long form: Will you want to ask, for example,

"How many FOX SHIPS are there?"

to get a count of the number of SHIPS
with NAM = FOX ?

Please type Y or N.

4. Questions Asked About Numeric Fields

Three different types of numeric fields are currently distinguished:

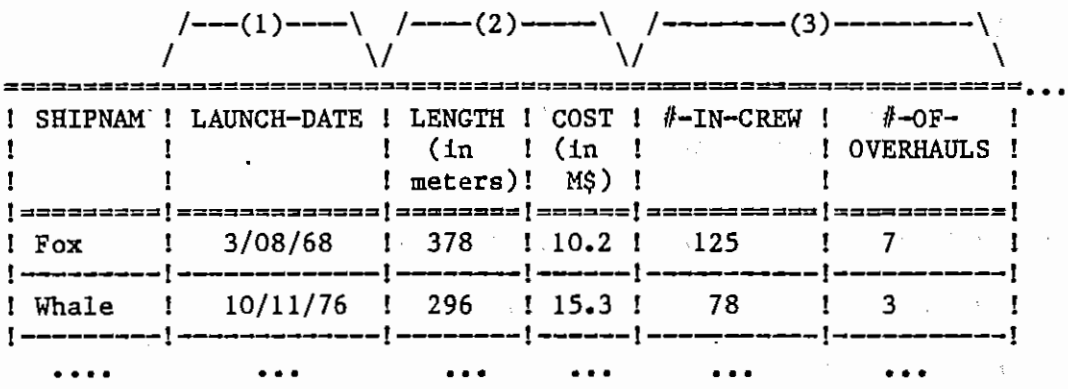
(1) Those whose values are dates (encoded in some particular way).

(2) Those whose entries are numeric values derived through some kind of measuring operation, corresponding to numbers of some kind of measurement units.

(3) Those whose values are "pure numbers," indicating the cardinality of some set of objects (concrete or abstract), rather than a set of measurement units.

To illustrate these distinctions, let us consider the following fields in a file about SHIPS:

SHIPS



These are examples of (1) date-valued, (2) measure-valued, and (3) "pure number-valued" fields.

The system first asks if the values in a numeric field are dates, then if they are measure numbers, and, finally, if they are "pure numbers."

a. Are Field Values Dates?

Short form: Are LGH's values dates (Y or N) ?

Long form: Are the values in this field, LGH, dates
(e.g., like July 11, 1947 or 7/11/47) ?

Please answer Y or N.

Current Limitation--If the user replies to this question in the affirmative, no furthur questions are asked by the present system. This has the following important consequences:

- The system has no knowledge as to what date-encoding conventions and date calculations the user obeys or requires.
- As a result, the system cannot determine whether one date is "before" or "after" another, and cannot answer questions that call for such a determination.

b. Are Field Values Numeric Measures?

Short form: Are LGH's values measurements (Y or N) ?

Long form: Are the numeric values in this field, LGH,
numbers that stand for measurements of something expressed
in some standard unit of measurement?

For example, numbers that record
* The length of something in feet or miles.
* The weight of something in pounds or tons.
* The cost of something in dollars.

Please answer Y or N.

c. Implicit Uniform Unit of Measure

Associated with the Field

Measure-number field values are associated with certain unit(s) of measurement (feet, tons, bushels, etc.). This association may be either

- Explicit or implicit. If the associated unit (or units) of measurement is marked in another field in the database, so that TEAM can look it up by scanning that other field, then it is "explicit." For example, in the following file:
  SHIPS

SHIP-NAME	CARGO	CARGO-QUANTITY	CARGO-UNITS
Fox	Oil	23,456	barrels
Adams	Wheat	987,716	bushels
Whale	Iron Ore	56,897	metric tons

  the CARGO-UNITS field explicitly specifies the units of measure associated with the numeric values in the CARGO-QUANTITY field.

  If the associated unit(s) of measurement does not appear in the database, but is understood by the database administrator, it is "implicit."

- Uniform or heterogeneous. If all the values in a measure-number field are associated with the same unit of measurement, then one has a uniform unit of measurement. If the values are associated with different units of measurement, as in the example immediately above, then one has a heterogeneous unit of measurement.

Current Limitation--The current system assumes, and is equipped to handle, only uniform and implicit units of measurement. So only the following question is asked:

Short form: What is the unit of measurement in which the values in this field, LGH, are expressed? (use singular; no abbreviation)

Long form: Are all of the numeric values in this field, LGH, measurements expressed in one and the same unit of measurement (e.g., foot, ton, etc.) ?

If not, type <CR>; if so, please type the unit of measurement in the singular (e.g., FOOT, not FEET; TON, not TONS).

For purposes of our examples, let us suppose that the user types "METER <cr>" in response to this question.

d. Plural Form of Unit of Measurement

Short Form: Plural of METER?
= [Default: METERS] Default accepted (Y,N)?

e. Synonyms for Unit of Measurement

Short Form: Synonym(s) for METER?

If the user specifies synonyms for the unit of measurement, the system will ask for the plural form of each synonym supplied.

f. Maximum and Minimum Values

This question must be asked in order to generate the correct database queries when the user requests a maximizing operation.
Short form: Minimum and maximum numeric values?

Long form: Please type the minimum and maximum numeric values that will be found in this field, LGH.

E.g., the minimum and maximum values for a numeric field specifying the speed of ships might be 0 and 68 (mi/h).

Please specify minimum and maximum numeric values.

g. Adjectives for Measure

Certain measures have adjectives associated with them (in all three forms--absolute, comparative, and superlative), as well as antonyms of those adjectives. (Fields with "pure number" values may also have similar associated adjectives: see Section 2.5.3.8 below.) TEAM can acquire knowledge of such adjectives to deal with possible user questions about measure fields. For example, in a file with subject SHIP and a measure LENGTH field, some plausible questions are

- "How LONG is the Fox (in meters)?"
- "What ships are LONGER than the Fox (or: than 500 feet)?"
- "What is the LONGEST ship?"
- "Is the Fox SHORTER than the Hoel?"
- "What is the SHORTEST ship?"

With some kinds of measure, such as BEAM, there is no associated adjective: one cannot ask *"How BEAMISH is the Fox?"--just "What is the BEAM of the Fox?"

The first question in the series below triggers its successors only if it receives a positive (nonnil) response.

(1) First, TEAM asks if there is an adjective corresponding to the field name of the measure field.

Short form:
An adjective associated with this arithmetic field, LGH ?

Long form: Is there a word wwww such that the question

"How wwww is the SHIP ?"

is equivalent to

"What is the LGH of the SHIP ?"

If not, type <cr>; if so, please type that word.

(Assume that the user's answer is LONG.)

(2) The system then asks for synonyms of the adjective provided.

Short form: Synonym(s) of LONG?

Long form: If there are synonyms of LONG, please list them.

If not, type <cr>; if so, type a list of words or multiwords.

(3) Now TEAM must check for "unusual" comparative and superlative forms of the adjective and its synonyms. This is done only if the adjective is a multiword. For example:

Comparative form of (ASW CAPABLE)?
= [Default: (MORE ASW CAPABLE)] Default accepted (Y,N)?

If the user types N, he must then type the correct form of the comparative. Similarly, the question for the superlative is the following:

Superlative form of (ASW CAPABLE)?
= [Default: (MOST ASW CAPABLE)] Default accepted (Y,N)?

(4) Now TEAM must ask the user if there are any antonyms of the adjectives just specified.

Short form: Antonyms for LONG?

Long form: Are there any antonyms of the word LONG?

For example, the antonym of LONG might be SHORT
the antonym of BROAD might be NARROW

If there are no antonyms, please type <cr>.
If there are such antonyms, please list them.

(5) Finally, if the user has specified antonyms for the adjective associated with the measure-valued field, TEAM must ask for "unusual" comparative and superlative forms of these antonyms.

h. Pure Number Field Values

This is the default case for numeric fields; if they have not been declared explicitly to contain dates or numbers, they are assumed to have "pure numbers" (i.e., numbers representing the cardinality of sets) as their values. Since there are no units associated with pure numbers,

TEAM asks only the following questions about such fields:

- The maximum and minimum values in the field (for the database). This question for pure-number fields is identical to the one about the maximum and minimum values of measure fields (see Section 2.4.4.6 above).
- Adjectives for the field name (e.g., "How HEAVILY STAFFED is the Fox?" for "What is the NUMBER-IN-CREW of the Fox?"), together with their synonyms and antonyms, plus the comparative/superlative forms of all of these. This series of questions is identical to the previous ones about adjectives for measurement fields (see Section 2.4.4.7 above).

5. Questions About Feature Fields

Feature fields are of two types.

Type I is a field with Boolean (i.e., only two possible and opposite) values--T/F, Y/N, POSITIVE/NEGATIVE, etc. Such a field represents a single property of the subject of a file and has values that indicate whether or not the property is possessed by the object a record is "about." An example in a file with subject SHIP would be a field ASW that shows whether a ship is capable of anti-submarine-warfare or a DAMAGED field that indicates whether the ship has incurred any damage.

SHIP

SHIP-NAME	ASW	DAMAGED	...
Fox	T	YES	
Hoel	F	YES	
Whale	T	NO	
...	...	...	

Note that a user could ask two different kinds of questions regarding such fields:

(1) Those exemplified by "What ships have an ASW of T?" (asked by a user knowledgeable about the actual contents of the database). Such questions are like "What ships have a LENGTH of 500 feet?" and rely on there being a binary predicate associated with the field name, and linking the field name with the subject of the file (a binary predicate like those associated with all other types of fields).

(2) Questions that seek to elicit what the Boolean feature field really represents (the presence or absence of a property, i.e., whether an object actually has that property) without even mentioning the actual database Boolean field values:

- "What ships are ASW-CAPABLE?" (here an adjective is associated with the field name--an adjective that applies to an object having the property).
- "What ships have (an) ASW CAPABILITY?" (here an abstract noun is associated with the field name--an abstract noun that names a quality an object that has the property is said to possess).
- "What ships are SUB-KILLERS?" (here a common, concrete noun "sub-killer" is associated with the field name--a common noun that refers to the class of objects that has the property).

To be able to respond to these various forms of questions, TEAM asks the user whether he will employ any such expressions.

Type II feature fields typically have more than two possible kinds of field values (although not necessarily); more importantly, they differ from Type I fields in that their values do not record whether or not an object possesses a property, but rather with what certainty the object is believed to possess the property.

An example of a non-Boolean feature field would be the INFECTIOUS

and METASTASIS fields of a file about cancer patients:

CANCER-PATIENTS

PATIENT-NAME	INFECTIOUS	METASTASIS	....
John Jones	probable	certain-positive	
Mary Kazlan	unlikely	possible	
Sam Snide	certain-negative	unlikely	
....	...	...	

Such Type II files share some of the characteristics of symbolic fields whose values are points on some scale (i.e., rank-ordered along some dimension). For instance, one could well imagine a user's asking such questions about the above file as

- "What patients have probable or more than probable metastasis?"
- "What patients are less than certainly infectious?"

and so on.

Current Limitation--The present system is not equipped to handle non-Boolean feature fields, nor does it ask any questions about them. All its questions concern Bool n ea re ields.

a. Positive Field Value

TEAM must know what the positive and the negative field value entries in a feature field are to handle questions referring to file subjects having the property represented in the field, such as "What ships have an ASW of T/F?" It first requests the positive field value.
Short form: Positive field value for ASW ?

Long form: Please enter the positive field value in
this (Boolean) feature field, ASW.

(For purposes of this example, we assume that the user answer is AA.)

b. Expressions for Querying the Positive

Boolean Field Value

These questions elicit adjectives, abstract nouns, and concrete nouns associated with the presence of a positive field value.

(1) Adjective(s).

Short form: Any adjective(s) associated with field value AA
in this ASW feature field ?

Long form: Are there any words wwww such that if you wished to
know whether

a SHIP has a ASW
of AA

you might then wish to ask the following kind of question?

"Is the SHIP wwww ?"

(For example, in a medical file about DISEASEs that has a feature field called CONTAGIOUS with a positive field value T, instead of asking "What diseases have an CONTAGIOUS of T?" one might want to ask "What diseases are INFECTIOUS?" so that "INFECTIOUS" would be the adjective associated with a positive value in the feature field CONTAGIOUS. Similarly, the adjective associated with the

negative field value F in the CONTAGIOUS field might be "UNINFECTIOUS".)

Please type <cr> or a word or words wwww after the prompt.
(By way of example, we assume that the user's reply is (ASW CAPABLE).)

(2) Abstract Noun(s).

Short form: Any abstract noun(s) associated with field value "AA" in this ASW feature field? (use singular)

Long form: Are there any words wwww such that if you wished to know whether

a SHIP has a ASW
of AA,

you might then wish to ask the following kind of question?

'Does the SHIP have wwww
(or: have a(n) wwww) ?'

(For example, in a botanical file about PLANTS that has a feature field called POISONOUS with a positive field value T, instead of asking "What plants have a POISONOUS of T?" one might want to ask "What plants have TOXICITY?" or "What plants have TOXICOGENIC CAPABILITY," so that TOXICITY and (TOXICOGENIC CAPABILITY) would be the abstract nouns associated with a positive field value T in the POISONOUS feature field. (Note that the singular form is required, e.g., TOXICITY rather than TOXICITIES.))

Please type <cr> or a word or words wwww after the prompt.
(For purposes of this example, we assume that the user's answer is (ASW CAPABILITY).)

(3) Concrete Noun(s).

Short form: Any common noun(s) associated with field value "AA" in this ASW feature field? (use singular)

Long form: Are there any words wwww such that if you wished to know whether

a SHIP has a ASW
of AA

then you might wish to ask the following kind of question?

'Is the SHIP a/an wwww ?

(For example, in a file about EMPLOYEEs, with a feature field called AMERICAN-CITIZEN whose positive field value was T and whose negative field value was F, instead of asking "Does Jones have an AMERICAN-CITIZEN of T?" one might want to ask "Is Jones a U.S. NATIONAL?" then (U.S. NATIONAL) would be a common noun associated with a positive field value in the AMERICAN-CITIZEN feature field. Similarly, the common noun ALIEN might be a common noun associated with the negative field value F in the feature field AMERICAN-CITIZEN. (Note that the singular is required, namely, ALIEN and not ALIENS.))

Please type <cr> or a word or words wwww after the prompt.
(For purposes of this example, we suppose that the user's response to this question is "SUB-KILLER <cr>".)

If the user specifies any abstract or concrete nouns in response to the above questions, the system will ask for their plural forms. For example:

Short form: Plural of (ASW CAPABILITY)?
= [Default: (ASW CAPABILITIES)] Default accepted (Y,N)? y

c. Negative Field Value

Short form: Negative field value for ASW ?

Long form: Please enter the negative field value in this (Boolean) feature field, ASW.

(For purposes of this example, we assume that the user's answer is XX.)

d. Expressions for Querying Negative Boolean Field Values

The system elicits from the user any adjectives, or any abstract and concrete nouns associated with a negative field value in the Boolean field.

The questions are exactly the same as for positive field values, if

the negative values are substituted for the positive ones above.

As before, the plural forms of nominal expressions provided by the user will be requested.

E. Acquiring New Verbs

Verbs added to a file may be used in querying that file. For example, explaining the verb "COMMAND" to TEAM enables one to ask not only "Who is the commander of the Fox?" but also "Who commands the Fox?".

Current Limitation--The TEAM system presently acquires only simple transitive verbs, that is, only verbs that link two entities (a subject and an object). Both the subject and the object of the verb must be in the same file; one of them must also be the subject of the file.

1. File for Verbs

After the user has typed VERBS to the top level, the system asks:
Name of table:

The user must then type the name of an existing table, e.g., *SHIP*.

2. Names of Verbs

The system then requests the verbs by asking the following question:
Short form: Verb(s) for this file SHIP? (use infinitive form)

Long form: When interrogating the TEAM system about the contents of this file *SHIP* you may wish to use certain verbs in your questions.

For example, given a file *SHIP,* with the subject SHIP and a file called CAPTAIN, which lists the captains of various ships described by the file, instead of (or as well as) asking questions like

"Who is the captain of the ship Intrepid?"

you might want to ask the equivalent question

"Who COMMANDS the ship Intrepid?"

For the TEAM system to understand such questions, you have to tell it about the existence and linguistic properties of the verb COMMAND.

The infinitive form of the verb(s) is required, for example, (to) COMMAND, (to) SWIM, etc., rather than, COMMANDS or COMMANDED, SWAM, SWUM, etc.

Please list the verb(s) you might use in asking questions about this file SHIP.

(Assume that the user's answer is OWN.)

The following questions are asked about each verb the user supplies.

3. Past Tense of the Verb

If it is assumed that the previous answer was OWN, the system asks the following:

Short form: Simple past tense of the verb OWN?
= [Default: OWNED] Default accepted (Y,N)? ?

If the user answers no, the correct past tense must be provided.

Assume that the user accepts the default, since "I OWNED a car a year ago" is what a native speaker of English would say.

4. Past Participle of the Verb

In the same way, the past participle is also obtained:

Short Form: Past participle of the verb OWN?
= [Default: OWNED] Default accepted (Y,N)?

Assume that the user accepts the default once more, as one can say "I have OWNED my apartment for three years."

5. What Kind of Verb

The following question attempts to discover what kind of verb it is.

Short form: Is OWN (1) intransitive, (2) transitive, or (3) ditransitive (Answer 1, 2, or 3)

Long form: Which of the following types of verb does OWN belong to? Is it a(n):

1) Intransitive verb
like "fall," which can only occur with a single subject noun--e.g., one can say "John falls" but not "John falls a staircase";

2) Transitive verb
like "catch" or "sing," which can occur with both a subject noun and a second object noun--e.g., one can say "Mary catchs the ball" and "Sam sings a hit tune";

3) Ditransitive verb
like "give," which can occur with a subject noun, an object noun, and a third, indirect object noun --e.g., one can say "Jane gives Sam a present."

Note that some transitive and ditransitive verbs can be used with less than their maximum number of nouns--e.g., one can say either "Sam sings a hit tune" or "Sam sings," either "Jane gives Sam a present" or "Jane gives generously."

You are requested to specify the maximum number of nouns the verb OWN COULD be associated with.

Type 1 (intransitive), 2 (transitive) or 3 (ditransitive).

When deciding the maximum number of nouns a verb can be used with, it is important to consider only noun phrases, not prepositional phrases. Thus, "fall" is intransitive; it can, of course, still appear with a prepositional phrase, as in "John fell down the staircase.".

Current Limitation--Here the user has only one choice: transitive.

The next two questions determine whether the subject or the object of the verb is simultaneously the subject of the file.

6. SUBJECT OF THE VERB

Short form:What sort of thing can OWN?

Long form:Among the following sorts of things with which the database is concerned,

(SHIP NAM ID NAT),

who or what can be said to OWN (e.g., to be referred to as subjects of the verb OWN?

For example, consider the verb COMMAND. In a file concerned with categories of things like SHIP, CAPTAIN, SPEED, etc., a CAPTAIN could be said to be the kind of thing which/who could COMMAND.

What sort(s) of things can OWN?

Assume that the answer is NAT (the name of a nation or the nationality field; that is, a nation can own a ship.

7. Object of the Verb

Simple query:What sort of thing can be 'OWNED?'
= [Default: SHIP] Default accepted (Y,N)?

Since the subject of the verb was not the object of the file, the

system in essence compelled the user to take the subject of the file--SHIP--as the object of the verb.

Otherwise, the two forms of queries are invoked:
Short form:What sort of thing can be OWNED?

Long form:Among the following sorts of things with which the database is concerned,

(EMP ID CAR),

who or what can be said to be OWNED (e.g., to be referred to as objects of the verb OWN?

For example, consider the verb COMMAND. In a file concerned with categories of things like SHIP, CAPTAIN, SPEED, etc., a SHIP could be said to be the kind of thing which/who could be COMMANDed.

What sort(s) of things can be OWNED?

If the subject of the file is not provided as either the subject or object of the verb, TEAM will object with the following message:

The answer is incorrect in the following way(s):

At the present stage of development of the TEAM system, only those transitive verbs whose subject and object form a <file-subject/file-field-name> pair can be handled.

Since the subject you specified for this verb was the file-subject, the object of this verb MUST be one of the file-field-names.

Please provide a response that is one of the file-field-names, or else back up (using ! as an answer) to change the subject of this verb. Try again.

8. The Verb Pattern

The next question makes sure that the transitive pattern is correct:

Short form: Can a(n) <NAT> OWN a(n) <SHIP> (Y or N) ?

Long form: Is it correct to say the following?

a(n) <NAT> OWNs a(n) <SHIP> , e.g.,
(the/a(n)) US OWNs (the/a(n)) (FOX).

(Answer Y or N)

9. Correspondence Between Verb and Noun

This final question about a transitive verb verifies the correspondence between the new verb and a noun already known. The query has a single form:

Short Form: If one says

a(n) <NAT> OWNs a(n) <SHIP> , e.g.,
(the/a(n)) US OWNs (the/a(n)) (FOX).

is that the same as saying the following?

(the/a(n)) NAT of (the/a(n)) (FOX) is (the/a(n)) US.

(Answer Y or N)

F. End of Acquisition

The user is given a last chance to review and change some of his answers. TEAM will ask the following before updating its internal structures:

You have at present completed specifying (and perhaps also modifying)

the information required by ACQUIRE concerning this file SHIP.

Do you wish to review and/or change any of the information you specified before this information is used to update TEAM's knowledge?
(Answer Y or N)

If the user answers no, the system will then proceed to update its internal lexicon, conceptual schema, and database schema. If the answer is yes, the system will go into its backup mode.

G. Backup Mechanism During Acquisition

As mentioned earlier, it is possible at any time to modify the answer to a previous query. This mechanism is invoked by typing the character ! instead of replying to a question. For example, let us assume that <carriage-return> was typed by mistake as an answer to the adjective question:

An adjective associated with this arithmetic field LGH ?

The user can then back up by typing ! after the next TEAM question*. The following exchange will occur:

So you (perhaps) want to change some previous answer...

Okay, type in:

some keyword(s)	involved in the previous question or previous answer which you wish to change, or
'!!'	to see all of the previous questions and answers in the order in which they occurred, or

*In a few cases, TEAM will interpret ! as an answer to a question... If this occurs, the user should keep answering questions with ! until TEAM understands it as a backup request.

```
     <cr>                to quit.

 > adjective

Pick one of the following...  Note: Questions to which you gave no
                                    response, as well as those to
                                    which you answered NIL, are
                                    shown as having been answered
                                    with NIL.

     Questions                                                  Answers
     =========                                                  =======

(Question #1 :)
AN ADJECTIVE ASSOCIATED WITH THIS ARITHMETIC FIELD
LGH ?
                                                                NIL

Type in the number corresponding to the question/answer you wish to
change, or <cr> to specify different keyword(s).
 > 1
```

The keyword feature was used in the above interaction. This could have been either a word mentioned in the question or one in the preceding answer. In this case, the keyword was adjective, which term was included in the question.

Instead of a keyword, !! could have been used. TEAM would have displayed all the questions with their answers and the number of the question to be modified.

The interaction is more complex when the user wants to modify a question, the reply to which is a list. In the following example, the user has previously specified (NAM ID) as the fields of the file SHIP and now wants to add to them (LGH ASW):

```
SHIP's fields ?
 = [Default (= previous answer): (NAM ID)] Default accepted (Y,N)?  N


The previous answer to this question was

                                    (NAM ID)

Please type:
          A--if you wish merely to add to the previous answer
          M--if you wish to modify the previous answer (and
             perhaps add elements to it as well)
          !--if you wish to back up to yet another question
          Q--if you wish to quit, leaving the previous answer
             unchanged

For purposes of this example, the user types A; the session then
resumes as follows:

Add to previous answer
Please type what you want added onto the end of the previous
answer--(NAM ID)

> LGH ASW

Your alterations have produced the following new answer:

                                    (NAM ID LGH ASW)

Are these alterations satisfactory (Y or N) ? YES
```

Note there are two different ways to get into this backup feature. First, as shown above, by typing ! while in the normal acquisition mode. Second, when ACQUIRE is typed at the top level, the user is first asked whether he wants to modify a previous table acquisition. If he answers yes, the user is in effect "backing up" to his previous answers. An example of such an interaction follows:

```
ACQUIRE

Do you want to add information to a table you have
already started to describe?
 > Y
```

```
Name of that table:  > ?

The table must be one of the following:
(SHIP)


Name of that table:  > SHIP

Do you wish to (further) modify your specifications of this file?
        SHIP
(Answer Y or N)
> Y


Do you wish to change or add to the list of previously specified
field names of this file SHIP?
(Answer Y or N)
> Y
```

(At this point, the user is exactly in the same situation as in the previous backup example; he can now add new fields to the file).

In contrast, the user is explicitly asking to modify a previous answer in the following interaction:

```
ACQUIRE


Do you want to add information to a table you have
already started to describe?
 > Y


Name of that table:  > SHIP

Do you wish to (further) modify your specifications of this file?
        SHIP
(Answer Y or N)
> Y


Do you wish to change or add to the list of previously specified
field names of this file SHIP in particular?
(Answer Y or N)
```

```
> N

So you (perhaps) want to change some previous answer...

Okay, type in:

     some keyword(s)     involved in the previous question or in the
                         previous answer which you wish to change, or

     '!!'                to see all of the previous questions and
                         answers in the order they occurred, or

     <cr>                to quit.

 > identity

Pick one of the following... Note: Questions to which you gave no
                                   response, as well as those to
                                   which you answered NIL, are
                                   shown as having NIL for an answer.

     Questions                                                  Answers
     =========                                                  =======

(Question #1 :)
SYNONYM(S) FOR ID?

(IDENTITY (IDENTITY CODE) (UNIT CODE) (UNIT IDENTITY CODE)
(UNIT IDENTIFICATION CODE))

(Question #2 :)
PLURAL OF IDENTITY?
                                                        IDENTITIES

(Question #3 :)
PLURAL OF (IDENTITY CODE)?
                                                (IDENTITY CODES)

(Question #4 :)
PLURAL OF (UNIT IDENTITY CODE)?
                                            (UNIT IDENTITY CODES)


Type in the number corresponding to the question/answer you wish to
change, or <cr> to specify a different keyword(s).
 > 1
==========================================================
```

```
Synonym(s) for ID?
 = [Default (= previous answer): (IDENTITY (IDENTITY CODE)
(UNIT CODE) (UNIT IDENTITY CODE) (UNIT IDENTIFICATION CODE)) ]
    Default accepted (Y,N)?  n


The previous answer to this question was


(IDENTITY (IDENTITY CODE) (UNIT CODE) (UNIT IDENTITY CODE)
(UNIT IDENTIFICATION CODE))

Please type:
          A--if you wish merely to add to the previous answer
          M--if you wish to modify the previous answer (and
             perhaps add elements to it as well)
          !--if you wish to back up to yet another question
          Q--if you wish to quit, leaving the previous answer
             unchanged

? add to previous answer
        ...
```

and the user now has a chance of adding new synonyms or modifying existing ones.

III OTHER MISCELLANEOUS FUNCTIONS IN TEAM

A. Database Entry

EDIT allows the user to enter data into his local database. It can be called only after at least one acquisition session has taken place. If only one file has been introduced, EDIT will immediately be ready to edit that file. Otherwise TEAM will ask which file should be updated at that point.

The database editor will then ask two questions if it is the first time it has been used during the session. The initial question refers to the size of the field names and values (the purpose of this is to optimize the use of the screen). The editor asks the following:

> If the length of the attribute names or the values in them exceed the default length of 14 characters, type that number (otherwise just <cr> for the default 14). >

Typically, all that is needed is a carriage return. However, in some cases a number (other than 14) may be appropriate, especially if the field names or values are usually long or short.

The following question has to do with the particular terminal being used:

```
        This editor will make use of the cursor control commands
    (if any) on your terminal. Please specify your terminal type:
         (1)  Datamedia 3000, 3025, 3045
         (2)  Datamedia 1500, 2500
         (3)  Heathkit 19
         (4)  None of the these; use a TTY editor
         (5)  Load a file containing the cursor control information.
    Please type 1, 2, 3, 4, or 5. >
```

For purposes of this example, let us assume that the user responds with 4 (which is always safe). TEAM will then go into its database editor.

The function EDIT allows the user to add, delete, and change the data in a table in a convenient manner.

Since it is difficult to give an adequate description here of the way the database editor works, it is recommended that the user familiarize himself with it by actually trying it out. The TTY editor, which is similar to the display editor but operates on any TTY-type terminal, is described below.

```
SPACE, >  = Move right     DEL, <  = Move left       ^  =  Move up
LINE-FEED = Move down      RETURN = Go to the front  of this line
{ = Go to the first record                 } = Go to the last record
CTRL/S = Find a value     CTRL/A = Add new records  CTRL/Q = quit
! = Reprint the attribute names          ctrl-K = Kill this record
? = Retype this header (if at :)
A number directly after a single : = Go to that record
<cr> after a :  =  Repeat the previous search

The number of :'s indicates which row of the record the
editor is on.  To change a value, move the cursor under that
value and type in the new value followed by either a space
or <cr>.

:   NAME            SEX             AGE             SSN       MASS


ADD RECORD. Finish with a <CTRL/Q>


1   ? BIL           ? MALE          ? 29        ? 123456      ? 74
2   ? KIM           ? FEMALE        ? 30        ? 654321      ? 66
3   ? SANDY         ? FEMALE        ? UNKNOWN   ? 111         ? 0
4   ? ^Q
3   SANDY           FEMALE          UNKNOWN         111       0
:                    MALE
3   SANDY           MALE            UNKNOWN         111       0
:{
1   BIL             MALE            29              123456    74.0
:^S
Attribute > SEX
Value > MALE
1   BIL             MALE            29              123456    74.0
:
3   SANDY           MALE            UNKNOWN         111       0
:^K
2   KIM             FEMALE          30              654321    66.0
:?
```

In the above example, the first few lines explains the editor commands to the user. When the user is not in the "add record" mode, he can type those commands to move the cursor as explained (for example, a > or space would move the cursor to the next field to the right). When

the cursor is in the proper place, the user can change the content of the field value by simply typing the new one, finishing with a carriage return. To go into "add record mode", the user types CTRL/A. He then types the field values separated by either a carriage return or a space (this is valid only for the TTY editor). The cursor will move to the next field position automatically. To get out of the "add record mode" the user types a CTRL/Q. Finally, to get out of the editor, the user types another CTRL/Q.

In the previous example, the user typed in three records. Next he positioned the cursor on the second field of the third record, and changed it from FEMALE to MALE. He then chose (via {) to look at the first record and, using CTRL/S, started a search for a record with the field value MALE for the attribute SEX. He deleted that record with a CTRL/K. Finally, the user typed ?. The system would then "refresh its display," i.e., show its commands and the current record again.

As already demonstrated, it is possible to modify values in the database. As a further illustration, if the user adds more fields to the file--such as HOME.TOWN in the following example-the session could proceed as follows:

[Note: the display of the editor commands does not reappear here.]

```
:    NAME            SEX            AGE            SSN       MASS
::   HOME.TOWN
1    KIM             FEMALE         30             654321    66.0
: <LINE-FEED>
1    NIL
::   PAO
1    PAO
::
2    BIL             MALE           29             123456    74.0
: <LINE-FEED>
2    NIL
::   (MENLO PARK)
2    (MENLO PARK)
:: ^Q
```

This example shows in particular the role of the : in the editor for handling cases in which a record cannot fit on one line of text.

Current Limitations--Important!

(1) All values entered into fields must be words, not multiwords. Thus if you wish to put in LOS ANGELES as the name of a particular ship, you must insert a "-" to make this phrase a single word, "LOS-ANGELES".

(2) Identical values should not be entered into two different symbolic fields. For example, consider the following comparison:

FILE-1 (incorrect)

NAME	RADIO-CALL-SIGN
FOX	FOX
WHALE	WHL
...	

FILE-2 (correct)

NAME	RADIO-CALL-SIGN
FOX	FX
WHALE	WHL
...	

FILE-1 would cause problems for TEAM, because the value "FOX" appears in two different symbolic fields.*

*This limitation will be removed in the very near future.

B. User Profile

Another TEAM top-level function is PROFILE. It is invoked by typing PROFILE to the TEAM top-level prompt as follows:

PROFILE

The system asks a series of questions to determine how best to interact with the user. Those questions are

Short form: Do you want to load your user profile? ?

Long form: TEAM has two alternative ways of setting your user profile:
1. By loading a file that contains it (and has been saved in a previous session)

2. By asking a series of questions.

Do you want to load your user profile?

The following questions are asked if the response is no (typical answers appear here in boldface):
Do you use a Datamedia 3000 (Y,N) ? > N

Do you want your queries to access a database? (If you answer NO, only natural-language parsing will be performed) > Y

Do you want to access an external database? (If you answer NO, a local TEAM database will be accessed.) > N

Do you want to see the logical forms generated by the natural language parsing component? > N

Do you want to save the above information in a user profile? Y

Name of file where user profile will be saved: PROFILE

At this point, the user has his profile saved in a file (here the file name is PROFILE). He can subsequently answer those same questions

implicitly by interacting as follows:

PROFILE

Do you want to load your user profile? Y

Name of the file where your user profile is saved: PROFILE

C. Saving and Loading the Schemas

After an acquisition has been performed, it is important to save the result of the session. This is done via the top-level command: SAVE. The interaction proceeds as follows:

SAVE

Print main name of files where internal data structures
are to be saved: SHIP

The file name must be a single word. There are actually three files created. In this case, they are: SHIP.ACQ, SHIP.NLF, and SHIP.LANG. Each has specific relevance to one portion of the system.

To load those files subsequently, the user invokes the TEAM top-level command: LOAD. An example of such an interaction follows:
LOAD

Print main name of files to be loaded: SHIP

D. Miscellaneous Commands

Two more commands remain.

The first one, LOCATIONS, allows a user to specify where his database is located--when it is not the local database.

Current Limitation--Because LOCATION is currently not implemented

and is not relevant to the NPS experiment, its description is omitted.

The final command is QUIT, which does just that. It is typically the last command of a session. Before QUIT is issued, an appropriate SAVE should be performed:

```
QUIT
```

IV QUERYING DATABASES

TEAM makes very strong assumptions about what you type to it. In particular, it assumes that database queries are relevant to the database it is designed to access. Questions about kinds of information not contained in the database (for example, sensor ranges in the Bluefile) normally cannot be interpreted. Questions about ships, places, or facts not in the database will result in either failures to parse or incomplete answers. For example, the question,

IS THE MARSEILLES A US SHIP?

cannot be parsed, because Marseilles is not in the database and hence not in TEAM's vocabulary. The question,

WHO COMMANDS THE SHIPS?

will not give information about any Soviet ships that might be retrieved, because no information on their commanders is available.

Questions can be asked about the entities in the database. Examples of simple restriction-type questions are

```
LIST THE LOS-ANGELES CLASS SHIPS
WHAT SHIPS ARE LONGER THAN THE FOX?
WHAT IS THE LONGEST SHIP?
WHAT IS THE LONGEST SHIP WHOSE FUELSTATUS IS ZERO?
WHAT SHIPS ARE FASTER THAN 20 KNOTS?
WHAT SHIPS ARE ASW CAPABLE?
WHAT SHIPS ARE CRUISERS?
```

However, most questions typically asked of a database are concerned

with the current values of attributes that are explicitly stored. TEAM provides many formats for specifying such questions. The simplest forms ask for the stored attributes. For example:

WHO IS THE COMMANDER OF THE FOX?
WHO COMMANDS THE FOX?
WHAT IS THE RADIO CALL SIGN OF THE FOX?
WHAT IS THE POSITION OF THE HOEL?
WHAT IS THE BEAM OF THE ASW CAPABLE SHIPS?

Others may use adjectives:

HOW LONG IS THE FOX?

Still others may compare values:

IS THE FOX LONGER THAN THE ENTERPRISE?

We hope the user will experiment with all these and other forms.

Current Limitations:

- TEAM cannot answer questions that rely on intra- or interfile linkages. An example of an intrafile linkage between fields is the situation in which the user has a CARGO-UNITS field specifying the units in which another CARGO-QUANTITY field is expressed (see Section 2.4.3.1). An example of an inter-file linkage is one in which the user has two files, a SHIPFILE and a SHIPTYPEFILE:

SHIPFILE

NAME	TYPE
FOX	DDG
HOEL	DDG
WHALE	CVA

SHIPTYPEFILE

TYPE	LENGTH
DDG	515
CVA	1029

Each ship contained in these files "inherits" the length of the type it belongs to, and knowledge of the interrelationships of the two separate files is required before the question "What is the length of the Fox?" can be answered.

- The system cannot answer questions that presuppose

knowledge of encoding conventions for dates and of the related calculation procedures. For example, it cannot determine "What ships have a LAUNCH-DATE earlier than 3/9/78?"

- Questions containing conjunctions (such as "and") are not interpretable. Thus, one cannot ask "What ships are ASW-capable AND are longer than the Fox?". Equivalent questions that avoid the use of conjunctions are often answerable, however, such as "What ASW-capable ships are longer than the Fox?".
- Questions involving negation ("not") may be uninterpretable. Such questions have not yet been tested. Thus "What ships are NOT ASW-capable?" may not be successfully processed.

REFERENCES

1. W. Teitelman, INTERLISP Reference Manual, XEROX Palo Alto Research Center, Palo Alto, California, 1974.