Introducing ASK, A Simple Knowledgeable System

Bozena H. Thompson
Frederick B. Thompson

Computer Science
California Institute of Technology

5054:TM:82
INTRODUCING ASK, A SIMPLE KNOWLEDGEABLE SYSTEM

Bozena H. Thompson
Frederick B. Thompson

5054:TM:82

Submitted to

Conference on Applied Natural Language Processing
Santa Monica, California
February 1-3, 1983

California Institute of Technology
Pasadena, California 91125

November 1982
ABSTRACT

ASK, A Simple Knowledgeable System, is a total system for the structuring, manipulation and communication of information. It is a simple system in the sense that its development concentrated on clean engineering solutions to what could be done now with good response times. The user interface is a limited dialect of English. In contrast to expert systems, in which experts build the knowledge base and users make use of this expert knowledge, ASK is aimed at the user who wishes to create, test, modify, extend and make use of his own knowledge base. It is a system for a research team, a management or military staff, or a business office.

This paper is designed to give you a feel for the general performance of the ASK System and overview of its operational capabilities. To this end, the movie you see will continue throughout the talk. Indeed, the talk itself is a commentary on this background movie. The movie is bona fide and in real time, it is of the ASK System in action. (Many of the illustrations from the movie are reproduced in the written paper.)

I. ASK AS A DATABASE SYSTEM

A. Examples of ASK English

To introduce you to ASK, we will start out with a few examples of queries of a simple data base concerning ships. The uninitiated user may wish simply to ask:

>How many ships are there? 7
>What is known about ships? some are in the following classes:
  Navy
  freighter
  old
  tanker
  all have the following attributes:
  destination
  home port
  some have the following attributes:
  cargo
  all have the following number attributes:
  age
  some have the following number attributes:
  speed
  length
  beam

>What are the destinations and home port of each ship.
ship   destination   home port
Ubu    New York   Naples
       Tokyo      ---
Maru   Oslo       Tokyo
Kittyhawk   Naples   Boston
           London     ---
Alamo  London     London
       New York   ---
North Star London     New York
Nimitz London     Norfolk
Saratoga unknown Norfolk
>What cities are the home ports of ships whose destination is London? Boston
London
New York
Norfolk
>Are there ships that do not have a cargo? yes
>What is the number of New York ships? There are 2 answers:
  (1) New York (destination) ships 2
  (2) New York (home port) ships 1
>How many ships are there with length greater than 600 feet?
Spelling correction: "lenngth" to "length" 4
>What ships that carry wheat go to London or Oslo? ships that carry wheat
London
Maru
Oslo
Alamo
>Does the Maru carry wheat and go to London? yes

B. The ASK Data Structures

Although in the terminology of data base theory, ASK can be considered as an "entity-relation" system, ASK retains its information in records which are interlinked in a semantic net. One reason we refer to ASK as simple is because it uses only a few kinds of nodes in its semantic net, namely:
and the obvious corresponding arcs. We speak of this as the COAR structure. Attributes are single valued, e.g., "father", "home port", "title"; relations may be multiple valued, e.g., "child", "cargo", "author". The difference between attributes and relations can be seen in the following protocol.

>What is the cargo and home port of the Maru?
cargo home port
wheat London
>The home port of Maru is Boston.
London has been replaced by Boston as the home port of Maru.
>What is the cargo and home port of the Maru?
cargo home port
wheat Boston
coil ---

C. Extending and Modifying the Data

1. Definitions

To make such a system more knowledgeable, one needs to be able to add definitions that embody interrelationships among the basic classes, objects, attributes and relations of the data. The simplest form of definition is synonym:

>definition:tub:old ship
Defined.

Although this form of definition allows one to introduce abbreviations and many forms of jargon, more extensive forms of definition are desirable. Here are three illustrations using the same "ship" file as above. In the third definition, note the use of quotes to create local "variables".

>definition:area:length * beam
Defined.

>definition:long:paper whose number of pages exceeds 49
Defined.
>definition:long:book whose number of pages exceeds 800
Defined.
>What AI bibliography items are long? There are 2 answers:
(1) long:paper whose number of pages exceeds 49
Physical Symbol Systems
A General Syntactic Processor
(2) long:book whose number of pages exceeds 800
Human Problem Solving
>What long books were written in 1972?
long:book whose number of pages exceeds 800
Human Problem Solving

Family relationships make for a good illustration of definitions; we switch to a small family relationship context.

>What are attributes?
individual/individual attributes:
spouse
>What are relations?
individual/individual relations:
parent
>What are classes?
individual classes:
male
female
>What are definitions?
definition:mother:female parent
definition:father:male parent
definition:child:converse of parent
definition:sibling:child of parent but not oneself
definition:cousin:child of sibling of parent
>What long books were written in 1972?
long:book whose number of pages exceeds 800
Human Problem Solving

2. Verbs

Most verbs embody knowledge specific to the application in which they are used, the exceptions being the copula verbs. Therefore the only verbs initially known to the ASK System are "to be" and "to have". The user can add new verbs by paraphrase.

>verb:ships "go" to New York:destination of ships is New York
Defined.
>verb:ships "carry" coal from London to Boston:ships have coal as cargo, have London as home port and go to Boston
Defined.
>Each old ship carries what cargo to each port?
old ship port cargo
Ubu New York oil
Alamo London wheat

The notion of what is long may be quite different in another context, say in the context of bibliography of artificial intelligence literature.
3. Anaphora: Pronouns and Ellipses

In practical systems for experts, abbreviated forms of addressing the computer are common. Thus the ability to handle pronominal and elliptical constructions are of considerable importance. Although there has been progress in the last few years in the linguistic understanding of these constructions, many difficulties remain. However, building on the work that has been accomplished, many of these constructions can be handled by the ASK System. In order to avoid misleading the user when the computational algorithm does not make the correct interpretation, echo is used to inform the user of the interpretation that has been taken.

>Is there a ship whose destination is unknown?
yes
>What is it?
What is it [ship whose destination is unknown]?
Saratoga
>Is its cargo wheat or coal?
Is its [ship whose destination is unknown] cargo wheat or coal?
wheat no
coal yes
>Does the Alamo's length exceed that of the Ubu?
yes
>What about the Northern Star?
[the Northern Star] does the Alamo's length exceed that of the Northern Star?
no
>Is its cargo wheat?
Is its [Northern Star] cargo wheat?
no
>Is it that of the Ubu?
Is it [Northern Star cargo] that [cargo] of the Ubu?
yes
>Wheat is carried by how many ships?
2
>Coal and oil?
coal and oil is carried by how many ships?
coal 1
oil 2
>by the Ubu?
coal and oil is carried by the Ubu?
coal no
oil yes

4. Extending the Vocabulary and Data Base

New objects, classes, attributes and relations may be added at will. This will be illustrated using the AI Bibliography.

>What is carried by the Alamo?
wheat
coal
>Wheat is carried to London from what ports?
New York
>What cities does the Alamo carry wheat to?
London

>What is it?
What is it [ship whose destination is unknown]?
Saratoga
>Is its cargo wheat or coal?
Is its [ship whose destination is unknown] cargo wheat or coal?
wheat no
coal yes
>Does the Alamo's length exceed that of the Ubu?
yes
>What about the Northern Star?
[the Northern Star] does the Alamo's length exceed that of the Northern Star?
no
>Is its cargo wheat?
Is its [Northern Star] cargo wheat?
no
>Is it that of the Ubu?
Is it [Northern Star cargo] that [cargo] of the Ubu?
yes
>Wheat is carried by how many ships?
2
>Coal and oil?
coal and oil is carried by how many ships?
coal 1
oil 2
>by the Ubu?
coal and oil is carried by the Ubu?
coal no
oil yes

>What AI Bibliography items are about knowledge representation?
Scripts, Plans, Goals and Understanding
A Framework for Representing Knowledge

D. Inheritance

Classes, as well as objects, may be the arguments or values of attributes and relations. This raises the question of inheritance of properties and attributes. The answer is to be
found in the following protocol, using the attribute of color; inheritance for relations is handled in similar ways.

>class: racing sloop, sports boat, Norwegian

The following classes have been added:
- racing sloop
- sports boat
- Norwegian

>attribute: color

The attribute color has been added.

>individuals: white, blue, black, gray

The following individuals have been added:
- white
- blue
- black
- gray

>Saratoga is a racing sloop.

Saratoga has been added to racing sloop.

>Racing sloops are sports boats.

Racing sloops have been added to sports boats.

>Alamo is industrial and Norwegian.

Alamo has been added to industrial Norwegian.

>The color of sports boats is white.

White has been added as color of sports boats.

>Maru's color is gray.

Gray has been added as color of Maru.

>Black is the color of industrial.

Black has been added as the color of industrial.

>Blue is the color of Norwegian.

Blue has been added as the color of Norwegian.

>What is the color of each ship?

There are 3 answers:
(1) ship color
   Maru gray
   Saratoga white
(2) Alamo as industrial
   ship color
   Alamo black
(3) Alamo as Norwegian
   ship color
   Alamo blue

II. INTEGRATION OF MULTIPLE OBJECT TYPES

A. Extension of COAR to Multiple Object Types

So far we have illustrated ASK capabilities using only two types of objects:
- individuals, e.g., "John Jones", "Maru"
- numbers, e.g., "34.6 feet", "length of Maru", "number of ships"

ASK has been designed, however, to facilitate many kinds of objects. This is a capability orthogonal to the simple COAR structure in that for any types of objects there may be corresponding classes, attributes and relations.

B. An Example: Texts and Text Files

We will illustrate this multiple object type capability with the additional object type: text. Once this new object type was added (together with procedures to manipulate texts, i.e., a "word processor") then text classes, individual/text and text/individual attributes and relations were immediately available. It was a small task to add an electronic mail system to ASK; all that was required was an addition to the authorization procedure that assigned to each newly authorized person a new text class as his/her mail box.

>Create the text/text relation: comment

The text/text relation comment has been added.

>File the Budget Proposal in budget correspondence.

Budget Proposal has been added to the text class budget correspondence.

>Who have I received mail from?

John Smith
Peter Jackson
III. More General Aspects of the ASK System

A. Response Times

The movie, which accompanied the oral presentation of this paper, demonstrated that the response time, i.e., the time between completion of the typing of the input by the user to the appearance of the response on the terminal, is very good. But the data bases used in the illustrations have been small, toy data bases. The following table gives average response times for a few cases using larger data bases. The query used for this illustration is:

>What are the destinations of tankers?

The response time is rather insensitive to the total number of individuals, classes, attributes and relations in the data base, depending primarily on the size of the relation (destination) and its argument (tankers). Suppose that there are m tankers in the data base and that n individuals have destinations, i.e., the size of the destination relation is n. The table gives time in seconds.

<table>
<thead>
<tr>
<th>no. of tankers</th>
<th>1</th>
<th>10</th>
<th>100</th>
<th>1000</th>
<th>2500</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>100</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>1000</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>2500</td>
<td>5</td>
<td>6</td>
<td>9</td>
<td>22</td>
<td>38</td>
</tr>
</tbody>
</table>

Response Time in Seconds for:

>What are the destinations of tankers?

B. The Concept of a User Context and the Basing Operation

In the terminology of ASK, a user "Context" is a knowledge base together with the vocabulary and definitions that go with it. A given user will usually have several Contexts for various purposes, some of which may be the small "Ships" Context, a (truncated) bibliography of Artificial Intelligence literature and an administrative Context concerning budget matters.

When one initiates a session with the ASK System, one is initially in the Command Context:

>Welcome to ASK
Please identify yourself.

Fred
>Pass word:
You have mail.

Fred is in COMMAND, proceed.

At this point, you can list the Directory of Contexts available to you, create or delete Contexts, authorize others to use Contexts which you have created, and enter any of the Contexts in
A new Context is created by basing it on an already existing one. Consider a user who has been authorized for basing on the AI Bibliography Context illustrated above and who wants to build a wider bibliography Context (adding new information -- vocabulary, data and definitions), however, without disturbing the old one. To do so, all s/he needs to do is select a new name, say CS Bibliography, and type:

> exit
  You are in COMMAND, proceed.
> Base CS Bibliography on AI Bibliography
  The new context CS Bibliography has been created based on AI Bibliography
  The result of this basing action is a new Context. Upon entering this new Context:

> Enter CS Bibliography
  You are in CS Bibliography, proceed.
  one can make additions:

> individuals: An Introduction to Database Systems, C. J. Date
  The following individuals have been added:
  An Introduction to Database Systems
  C. J. Date
> The author of An Introduction to Database Systems is C. J. Date.
  C. J. Date has been added as author of An Introduction to Database Systems.
> Keyword of An Introduction to Database Systems is database.
  database has been added as keyword of An Introduction to Database Systems.
> Who wrote what about databases?
  author
  D. L. Waltz Natural Language Access to a Large Data Base
  L. Harris Experience with ROBOT
  exit to CS Bibliography,
  You are in CS Bibliography, proceed.
> Who wrote what about databases?
  author
  D. L. Waltz Natural Language Access to a Large Data Base
  C. J. Date An Introduction to Database Systems
  L. Harris Experience with ROBOT

Several Contexts can be based on a given one, and one Context can be based on several, thus a hierarchical structure of Contexts can be realized. All Contexts are directly or indirectly based upon the BASE Context, which contains the function words and grammar of the ASK dialect of English, the mathematical and statistical capabilities, and the word processor.

C. Transportability

It is easy and fast to apply ASK to a new domain, given that a database for this new domain is available in machine readable form. The vehicle is the ASK dialogue-driven Bulk Data Input capability, which can be called upon to build an existing database into one's Context. The result not only integrates this new data with that already in the Context and under the ASK dialect of English, but in many circumstances will make the use of this data more responsive to users' needs.

The Bulk Data Input Dialogue prompts the user for necessary information to (1) establish the physical structure of the database file to be included, (2) add necessary classes and attributes as needed for the new data entries. The user also indicates, using English constructions, the informational relationships among the fields in the physical records of the database file that s/he wishes carried over to the ASK Context.

IV. DIALOGUES IN ASK

Some have raised the question whether natural language is always the most desirable medium for a user to communicate with the computer. Expert systems, for example, have tended to use computer guided dialogues. One simple form such a dialogue
might take is illustrated by the following which a new entry is added to the AI Bibliography:

>New bibliography item
>Add to what bibliography? AI Bibliography
>Title: Natural Language Processing
>Author: Harry Tennant
>Keyword: natural language
>Keyword: syntax processing
>Keyword: speech acts

Natural Language Processing has been added to AI Bibliography.
>Title: The "new bibliography item" dialogue is completed.
>What AI Bibliography items were written by Harry Tennant?
Experience with the Evaluation of Natural Language Question Answerers
Natural Language Processing

Other alternative media for user/system communication are menu boards, selection arrays and query by example. Many other cryptic ways to communicate user needs to a knowledgeable system can be thought of; often the most useful means will be highly specific to the particular application. For example, in positioning cargo in the hold of a ship, one would like to be able to display the particular cargo space, showing its current cargo, and call for and move into place other items that are to be included.

In the past, enabling the system to respond more intelligently to the user's needs required the provision of elaborate programs since the user's tasks may be quite involved, with complex decision structures. The introduction of terse, effective communication has incurred long delays and thus the changing needs of a user had little chance of being met. In the ASK System, the users themselves can provide this knowledge. They can instruct the system on how to elicit the necessary information and how to complete the required task. This ASK capability is quite facile, opening the way for its ubiquitous use in extending the knowledgeable responsiveness of the computer to user's immediate needs. ASK includes two system-guided dialogues, similar to the Bulk Data Input dialogue by which users can instruct the System on how to be more responsive to their needs.

A. Forms Designing Dialogue

The Form is an efficient means of communication with which we are all familiar. A number of computer systems include a Forms package. For most of these, however, filling in a Form results only in a document; the Form does not constitute a medium for interacting with the knowledge base or controlling the actions of the system. The ASK Forms capability enlarges the roles and ways in which Forms can be used as a medium for user interaction. As the user fills in the fields of a Form, the System can make use of the information being supplied to (1) check its consistency with the data already in the knowledge base and, if necessary, respond with a diagnostic, (2) fill in other fields with data developed from the knowledge base, (3) extend the knowledge base, adding to the vocabulary and adding or changing the data itself, (4) file the completed form in prescribed files or in those indicated by the user and also mail it to a specified distribution list through the electronic mail subsystem. Since the Form processing can check consistency and modify the knowledge base, Forms can be used to facilitate data input. Since Form processing can fill fields in the Form, the forms capability includes the functions of a report generator. Letters and memos can be written as special cases of Form filling, automatically adding dates, addresses, etc. and filing and dispatching the result.

It must be easy and natural to add new Forms, if they are to be a convenient tool. That is the function of the Forms Designing Dialogue. Much like the Bulk Data Input Dialogue, the Forms Designing Dialogue holds a dialogue with the the user through which s/he can specify the fields of the Form itself and the processing of the above kinds to be automatically accomplished at the time the Form is filled in. Here is a simple example of a form that was designed using the Forms Designing Dialogue.

>What is the home port and commander of each old ship?
There are 2 answers:
(1) There is no commander.
(2) old ship home port
Ubu Naples
Alamo London
>Who is John Smith?
The following words are not in the vocabulary:
John Smith
>Inventory of wheat and corn oil?
List of wheat and corn oil inventory
wheat 86.7
corn oil 123400.
Note that the home port of the Alamo is London and that it does not have a commander, further that John Smith is not known to the System.
>Fill shipping

(For the purposes of the published paper, in contrast to the film shown at the presentation of the paper, only the initial and final copies of the form are given, underlines indicate fields filled in by the "user", the other fields automatically being filled by the System.)

(befor)

<table>
<thead>
<tr>
<th>ship:</th>
<th>quantity</th>
<th>item price</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>commander:</td>
<td>$</td>
<td>$</td>
<td></td>
</tr>
</tbody>
</table>
Shipping Form

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>wheat</td>
<td>35.75</td>
</tr>
<tr>
<td>500</td>
<td>corn oil</td>
<td>2.50</td>
</tr>
<tr>
<td>107.25</td>
<td>$</td>
<td>1250.00</td>
</tr>
</tbody>
</table>

commander: John Smith

Shipping List for Alamo has been filed in Shipping Invoice File.
Shipping List for Alamo has been mailed to Jones.

mail to:
Fill shipping has been completed.

> List the home port and commander of each old ship.
  > Old ship home port commander
  > Ubu Naples John Smith
  > Alamo London John Smith

> Inventory of wheat and corn oil?
  > Wheat and corn oil inventory
  > Wheat 83.7
  > Corn oil 122900.

> What is in the Shipping Invoice File?
  > Shipping List for Alamo

V. ACKNOWLEDGEMENTS AND CURRENT STATUS

The three System guided dialogues, Bulk Data Input, Dialogue Designing Dialogue and Forms Designing Dialogue, are from the doctoral dissertation of Tai-Ping Ho. The aspects of ASK concerning basing one Context on another are from the doctoral dissertation of Kwang-I Yu. The methods for handling anaphora, fragments and correction of inputs are from the doctoral dissertation of David Trawick.

ASK is implemented on the Hewlett Packard HP9836 desktop computer. To handle Contexts of reasonable size, one needs a hard disk. An HP9836 with an HP9725 disk was used in the illustrations in this paper. Our work is supported by the Hewlett Packard Corporation, Desktop Computer Division.