Q&A: Already a Success?

Gary G. Hendrix
Symantec Corporation
Cupertino, CA 95014 USA

When Prof. Wolfgang Wahlster (the organizer of this COLING-86 panel on "Natural Language Interfaces: Ready for Commercial Success?") sent out invitations to panelists, he stated that his goals were "to evaluate three natural language interfaces which were introduced to the commercial market in 1985 and to relate them to current research in computational linguistics." For comparison, he has asked each of us panelists to answer a standard set of questions. These I will answer, but first let me set the stage by answering two logically prior queries.

00) What is your commercial system?

Our system, called **Q&A**, was introduced by Symantec in September of 1985. Q&A is an integrated package for the IBM PC/XT/AT and compatibles. Its major modules are a file management system, a report generator, a word processor (which was used to compose this document), a spelling checker and "the intelligent assistant" or "IA." The IA lets users manipulate databases and produce reports by issuing commands in English.

Note that Q&A is NOT sold as a natural language interface per se, but as an integrated business-productivity tool.

0) Why did you give your natural language subsystem such a pompous name?

Our marketing department made me do it.

The Wahlster Questions

1) How successful is your commercial system?

How many copies have been sold? Q&A's monthly sales are measured in thousands. SoftSel, the major US distributor of software for microcomputers, publishes a "hot list" every two weeks ranking its best selling products. Q&A has been in the top 10 since February of this year, and has been as high as number 2. At the time of this writing, it is number 3, below only dBASE III and Lotus 123.

Are they being used on a regular basis? Hundreds of people use Q&A every day, and thousands use it frequently. Although we have not conducted a rigorous survey, straw polls indicate that about half the active Q&A users use the IA on a regular basis. (Many people use the product mostly for WP.)

Are there cost-benefit analyses? No, but with a $299 suggested retail price in the U.S. market it doesn't take much analysis to justify the cost. Since Q&A integrates modules for filing, reporting, writing, and spelling along with the IA, the pro rata price of the IA is only $60, which can often be recovered by answering a single question!

Has the system been empirically evaluated? Q&A has been reviewed by every major (and most minor) publications addressing the IBM PC market, and has received consistently high reviews. *Infoworld* gave Q&A a 5 disk rating --its highest. *PC Week* has called Q&A the "quintessential management tool." Q&A was the first software product ever featured in *Der Spiegel*; it took honors as the software "product of the year" in Australia; and it received an unprecedented 2-part review in the *New York Times*.

Q&A was also evaluated by the National Software Testing Laboratories, widely considered to provide the microcomputer industry's most objective testing, in a comprehensive survey of file management systems for IBM PCs. Results were published by *Software Digest* in February, 1986, in a 68-page report covering both evaluations and methodologies. Q&A received the highest overall evaluation ever given to any product in any category, and was tops in the critical areas of ease-of-use and ease-of-learning. In fact, Q&A was tops in all categories except speed, where it placed second against a system that keeps all its data in RAM rather than on disk.

How well do users adapt to system limitations? Users have been remarkably forgiving of natural language limitations, as measured by calls on our telephone support hot line. The chief complaint about the IA is not its language limitations, but its speed on floppy-based systems. (On an IBM AT with hard disk, performance is not an issue.)

What are the main areas of application? Among the most popular in a straw poll were: customer lists, personal records, simple inventories, expense reporting, bibliographic data, tax records, project records, job costing, employee records, enrollments, real estate, church records, prospect tracking.

2) What are the most important factors for success of a NL interface?

The key is to realize that users do not care about NL interfaces--they only want to get a job done. Thus the important thing is to help the user accomplish his task.

To help users do their jobs, Q&A provides both a NL interface and formal interfaces. Some customers like the formal interfaces so well they never bother with the IA. (This troubles me, but I take consolation in the facts that these users are in the minority, they appear to be happy, and we charge the same whether they use the IA or not.)

As for getting the job done through NL, good

linguistic coverage and robustness in parsing are clearly important. Yet I was amazed to discover that users are often content to find ANY English phrasing that will retrieve the data they need, and don't seem too worried if their most natural way of phrasing a request doesn't work.

For a generic database product, being able to port to new sets of data is a necessity. Q&A's IA is adapted to new data bases primarily through automatic means: The IA uses information from the database itself to extend its vocabulary. Users can also provide additional vocabulary manually through the TEACH module. This module is composed of eight subsystems for giving the IA "lessons" about words that will be used with the new database.

Far more human-factors effort went into the development and testing of TEACH than into any other part of Q&A. But the results are mixed. While, we believe we have produced by far the most easily ported NL system on the market, at least 10% of our customers contacted in a straw poll indicated that they were intimidated by the TEACH lessons, and therefore are shunning the IA in favor of the formal systems for manipulating the data.

There are no provisions in Q&A for modifying the underlying grammar rules. This is because Q&A has been designed as a product for end-users, not as a system-building tool for programmers or linguists. If some users are intimidated by even the simple vocabulary acquisition lessons of the Q&A TEACH module, it seems obvious that allowing users to modify the grammar would court disaster.

3) Isn't it too early for commercial NL systems?

Q&A sales reports seem to indicate otherwise.

4) Will customers lose interest in NL if overblown expectations cannot be fulfilled now?

In advance of Q&A's introduction, Symantec's whole management team was worried about inflated expectations, and a concerted effort was mounted to meet this challenge. Of course, we did all we could to make the NL capability robust. But we also undertook the task of educating the microcomputer industry infrastructure BEFORE introduction. This was done through countless interviews with industry analysts, financial people, authors, distributors, retailers, lecturers, the press, consultants, and executives from various hardware and software firms. Through their influence, these people help set the expectations of the market.

To my surprise and relief, most of these industry insiders seemed to expect less from NL than Q&A delivers. We are benefiting from the backlash to AI being oversold previously.

As for the customers themselves, they divide into two groups, with different expectations. Those who have not used computers before don't know what to expect. If anything, they think these "electronic brains" already use NL. So even if they are sometimes disappointed by the limitations of today's NL systems, the disappointment is less than it would be using conventional interfaces.

Those who already use computers either don't care about NL (because they take pride in having learned formal systems) or they are encouraged by how much CAN be done with NL. We have been pleased to find that many power users of micros choose the IA rather than Q&A's query-by-example system. These users appreciate the power of natural language.

There is, of course, great room for improvement in commercial NL systems. Early attempts like Q&A are useful in educating the market about NL's potential and in setting reasonable expectations for future products.

5) Is it just a tremendous amount of engineering to transform a research system into a commercial product or is it much more (e.g. new concepts, algorithms)?

The answer to this question depends upon what the "product" is. To turn a research system into a language development tool for programmers involves mostly straightforward engineering. (Yet such a tool could be important in the right niche markets.) Creating a viable end-user product for the mass market is another story. In this arena, product marketing (i.e., designing a product from the user's point of view to fill a real need) and craftsmanship are even more important than the enabling technology.

To produce Q&A, Symantec devoted only limited resources to the development of the NL analysis system per se. Far more time was devoted to human-factors and performance issues, which are not generally important in NL research systems. Considerable intellectual attention was focused on TEACH, on extracting NL related information from the database, on parse-time error handling, and on developing robust parsing methodologies. And all of these activities were dwarfed by the efforts to develop the database, report generator and word processor with their concomitant support software.

6) How will the market for NL interfaces evolve? (Gee, I wish I really knew!)

The fact that Q&A is so high on the sales charts is not likely to escape the attention of other developers of software for micro-computers. Because part of the program's success is clearly attributable to the IA, it is reasonable to expect other microcomputer players to begin incorporating NL in their products also.

Lotus Development Corp. recently bought GNP, the developer of a natural language interface for 123. Thus, the most visible player in the microcomputer software industry has announced its interest in the NL area. If

this product reaches the market bearing the Lotus brand name, natural language interfaces will have received a major endorsement.

Another factor is that personal computers may soon reach significantly lower price points. When high-resolution micros with a megabyte of memory and fast processors (80386 or 68000 family) reach the $1000 level, another market tier will open, attracting a new and larger group of people to computers. This group is distinct from the number-hungry group that made spread sheets popular. It will be more symbol-oriented, wanting word processing and access to on-line (or CD-ROM) data. NL interfaces can play a pivotal role in reaching this potential market.

I think NLP will be widely used to help do sophisticated spelling and grammar checking in support of word processors. In the short term, this application of NLP may be more important commercially than NL interfaces.

Speech input, as usual, is "just around the corner."

7) Should commercial systems be designed for mainframes, AI workstations or micros?

Obviously, microcomputers are too small for NL systems. So I advise all of you to concentrate on mainframes. Better, build your own special hardware to support NLP.

8) Is it a market advantage to be multilingual?

THE advantage of NL systems is understanding one's mother tongue. What good is an English NL system to me if I only speak German? I would be better off with a formal system!

Each major natural language defines a separate market. To reach all these markets requires a system for each tongue.

Symantec clearly recognizes the importance of translating Q&A into multiple languages. F&A, the German language version of Q&A, is already in an advanced state of development, and work has begun on other languages.

9) What research seems most important for improving the quality of commercial products?

In "Natural Language Processing: The Field in Perspective" (*Byte*, 9/81), Earl Sacerdoti and I sketched three types of NL systems. Q&A belongs to the most primitive of these three types. I do not believe that the technology, the hardware or the market is ready for the more advanced types of NL systems, but it is upon these types that our basic-research energies should be concentrated.

As for applied research, it would be useful to focus on database transportability. TEACH is clearly the weak link in Q&A.

Also, now that thousands of people have NL systems, our community desperately needs to discover how they are really being used. Symantec has a great deal of anecdotal evidence about the use of Q&A, but lacks the resources to do anything approaching a thorough study. (And besides, the study needs to be done by psychologists, not our crew of computer and language hackers.)

10) How do you assess the role of...

NL generation: After Q&A has interpreted a user's input, its proposed response is presented in English so that the user may be advised about what Q&A plans to do. We believe this feedback is essential, and that this and other applications give NL generation a bright future.

Cooperative response generation: This is an obvious next step--simple, cheap and useful.

Speech act recognition: Some steps in this direction have already been taken in Q&A (as in "I wish I knew the time"), but I believe we are a long way from having any viable commercial system that incorporate a meaningful theory of speech acts.

User modeling: Anything other than the most superficial modeling probably requires too much memory for today's micros.

Hybrid representation languages & new types of grammars: Those that are not just fads could have an impact in a few years, but I don't see them impacting the next generation of commercial products, at least not on microcomputers. Personally, I weary from seeing the same old stuff redone each year in a new programming language. I suppose some progress is made in this process, but it seems painfully slow. It would be better to focus more of our community's resources on natural language problems, rather than on polishing programming language techniques.

Multimedia communications: A big opportunity exists today for mixing pointing with NL input. There are obvious applications when dealing with land maps or other representations of space.

One More Question: What does it all mean?

The early indications are that the market is enthusiastically accepting Q&A and its NL component. Within a year we should know whether NL ability of the kind incorporated in Q&A is simply a passing fad or is perceived by the market as providing real value. If the latter is the case (and I believe it is), then the market will begin to demand NL capabilities in more and more products. Either way, there will be much exciting work to do.

Acknowledgment: Major contributions to the design/implementation of the IA were made by Dan Gordon, Brett Walter and Denis Coleman.