THE CRITICAL FEATURES OF INTERACTIVE QUERY SYSTEMS (LES CARACTERES CRUCIAUX DES SYSTEMES DE CONSULTATION INTERACTIFS)

Gerald X. Amey
Defence Research Board
Ottawa, Ontario

ABSTRACT

The principal factors affecting acceptability and cost effectiveness of interactive query systems are considered from the viewpoint of the user under the headings: query language, sign-on protocol, error recovery, accounting and indexing of information. Examples are given from experience with the retrospective search system of the Defence Research Board. It is contended that the value to the user of a system accessed by terminals is greatly enhanced if the same methods and query language can be used to tap the resources of external systems. The paper discusses system parameters where commonality is desirable. Technical problems associated with achieving a suitable level of compatibility between autonomous systems are examined. (Les principaux facteurs qui influent sur l'acceptation et l'efficacité par rapport aux frais d'exploitation des systèmes de consultation interactifs, sont étudiés du point de vue de l'utilisateur, sous les rubriques suivantes: langage de la consultation, mode d'entrée dans le système, reprise sur incident, comptabilité et indexation des données. Des exemples sont donnés à partir de l'utilisation du système de recherche rétrospective du Conseil de recherches pour la défense. On affirme que dans le cas de l'utilisateur, la valeur d'un système consulté par terminal s'accroît de beaucoup si les mêmes méthodes et le même langage peuvent servir à mettre à contribution les ressources de systèmes externes. L'étude traite des paramètres de systèmes qui présenteraient les traits communs désirés. On y considère aussi les difficultés techniques reliées à la réalisation d'un degré convenable de compatibilité entre des systèmes à fonctionnement autonome.)

CRITERIA FOR QUERY SYSTEMS

INTRODUCTION

In the information business, the interactive query systems now becoming available may appear to be the first step toward that promised land where anyone, anywhere, can locate any data he desires, almost instantaneously. What features would we ask of such systems? To what extent is it feasible to provide these features in today's systems?

CRITERIA FROM KATZ REPORT (Katz, 1969) (Science Council Report No. 6)

In the Science Council report which recommended a policy for scientific and technical dissemination of information, it was stressed that it must be possible to query a national system in English and French with equal facility, and that the effect of distance from source to user be minimized. We therefore have two basic criteria - bilingualism and parity of access cost for users irrespective of distance from the source. It was also underlined in this report that no country, and certainly no separate organization, could possibly assemble all the information which it may need, at a single location. Thus, a third criterion is the compatibility of a given system with other systems.

FUTURE SYSTEMS

The ideal system would probably be capable of voice input and provide graphical, textual, and sound outputs. Indeed, genuine progress has been made in this direction. Still, in the context of current technology, one must consider the normal mode of interaction with online query systems to rely primarily on keyboards for entry and display tubes or printers for outputs. In what follows, I shall discuss criteria applicable to conventional query systems of the DIALOG (Summit, 1967) type rather than idealized future types. Costs are not specifically treated since this subject has been thoroughly discussed (Krevitt, 1973) and it is not desired to get involved in convoluted cost-benefit controversies. Suffice it to say that, if the costs of using a system are visible and funds are provided by the user, he will be able to determine whether the cost of using a given system is reasonable.

THE USER VIEWPOINT

If we are to speak about critical features, it is important to name the class of individual to whom these features are critical. In this matter I must take the stance of ombudsman for the user. That is, I am trying to determine those features which are critical in achieving satisfaction for the user of a system. It is the

system manager who must solve the problems that arise in trying to meet these criteria. From the worm's eye view of the user and looking out into the universe of information accessible in one way or another, I ask what are the most harassing features of current information gathering methods? The first one, it seems to me, is that one must always work through an intermediary. Some examples of current practice will underline this point.

INDEXING CONSTRAINTS

Let us consider a user seeking material in a given subject area. If the material is classified according to some esoteric system favoured by librarians, the user must work through someone skilled in the particular classification. In a traditional library where classification is associated with shelving layout, he will at least get rapid feedback to aid his search by browsing in the stacks if this is permitted. Creative searching (as distinct from formal search-definition) absolutely requires the browse capability, or direct access of the user to the resource material.

ORDERING IDENTIFIED MATERIAL

The user may discover a book or journal article which he would like to see, through some system of selective dissemination of information (SDI) or other means. In many cases, the SDI note will not have pinpointed a convenient source. The individual must go to his librarian who will then use standard bibliographic aids to determine sources and order the material via usual administrative processes. If the book or article is not held locally, it may be several weeks before the user receives the material. This method, though cumbersome, slow and costly, still presupposes considerable library resources at the user's site.

These examples, I think, indicate a most desirable feature of a user-oriented system. The user should have the option of controlling all phases of search and acquisition, with a minimum of administrative overhead. Of course many users will still opt to work through traditional information middlemen.

SPECIAL PROBLEMS INTRINSIC TO MACHINE SYSTEMS

I shall now consider machine systems for retrospective searching based on terminals, and the additional problems they present. In the first place, there are formal procedures prior to signing on to the system. After acquiring a terminal, it is necessary to open an account with the firm providing the computer service, as also with the

communications company that controls the line between the terminal and the data bank. In general, separate contracts must be opened with all utilities involved, one with each computer centre and one for each different telephone line. It is immediately obvious that such a system imposes an almost insurmountable obstacle on all but the most determined would-be users of a system. If the user's needs could be neatly categorized in such a manner that he frequently requires the services of one, and only one system, such a barrier might be acceptable.

NON-HABITUAL USER OF A SYSTEM - ACCOUNTING PROBLEMS

Information users' needs, however, are ever extending in the direction of cross-disciplinary data. In addition, the policy shift in research effort from long-term fundamental to short-term applied research means that the area of a user's interests is constantly changing. As long as separate accounting for different systems is used, each such cross-disciplinary application would call for formalization in a special-purpose data bank. As I have pointed out elsewhere (Amey, 1971), flexible intercommunication between information systems can enable the formation at short notice and low cost of meta-systems (Amey, 1970) to serve missions which it is not desirable to institutionalize. I shall return to this later. From the user's point of view it is most desirable that access to all non-proprietary data banks be as easy as making a telephone call. Whether one wishes to speak with someone in his own home town, or someone provinces or continents away, the access method is the same and there is a minimum of administrative barriers. The problems of interfacing between different common carriers, varying rate-structures, and switching are not visible to the telephone user. In systems jargon, the service is transparent to the user. He pays one bill which includes routine service as well as long-distance and other charges. A common carrier system or utility that is operationally transparent to the user is an essential prerequisite for full development of information retrieval systems, if they are to move beyond the current rudimentary stage.

THE HANDSHAKE - LANGUAGE PROBLEMS

Even if the mechanisms for opening accounts have been streamlined by interposing a common carrier interface, one immediately runs into other problems as soon as one turns on the terminal. The protocol at signing-on is in general idiosyncratic. Each system has its own rigid requirements of format and terminology for entry to the system. After penetrating the system (one sometimes feels that crypt-analysis is needed to tune in to some systems) one is confronted with a query language that is unique to the particular system. Beyond this again, in many cases as in the second example above, lies the barrier presented

by the formal indexing language used in the classification of material. The above indicates, I think, critical areas crying out for solution if the user is to move into the wonderland where all the world's information is at one's fingertips (via a keyboard). I shall not speak in detail of the man-machine interface or of relevance and these technical factors which have been treated exhaustively elsewhere. (Amey, 1970; Meadows, 1970)

EXAMPLE: SOCRATES RETROSPECTIVE SEARCH SYSTEM OF DRB

I list below what appear to me to be the most critical parameters of interactive query systems, above and beyond those technical characteristics which are already treated fully in the literature. Examples of dialogue from the SOCRATES retrospective search system of the Defence Research Board are given to indicate one approach to satisfying these criteria. These are taken from the system manual (DSIS, 1973). The system is described fully elsewhere (Irvine, 1973; Smithers, 1974) but a thumbnail sketch follows to enable the reader to follow the example. See Table 1 for basic system parameters.

For security reasons the system is a hybrid (Currie, 1973) of micrographic and computer systems. There is a hierarchy of storage such that search parameters are stored in the computerized data base while full bibliographic data are stored in microfilm on cassettes and are searchable by a page-search reader. Full text of documents is available on microfiche. Thus a fully equipped station with both a computer terminal and microfilm readers enables a user to browse through the data at whatever level is chosen. The network communications, processing, and accounting are provided by the host-system of I.P. Sharp Associates Limited. The system is designed according to the precepts previously offered (Amey, 1970) to be user-oriented, whether the inquirer is a casual user (conversational mode) or a habitual user (command mode), anglophone or francophone. Communication costs are bundled in the computer connect-charge and are independent of location. Sign-on protocol is not shown for security reasons.

CRITICAL PARAMETERS OF INTERACTIVE QUERY SYSTEMS

access policy - parity of anglophone and francophone in querying communications independent of location

ease - minimum complexity of sign-on protocol - independent of data base - favourable to casual users

billing - bundled accounting covering communications, processing and data base use (internal and external)

compatibility - interface problems should not be visible to user

CRITERIA FOR QUERY SYSTEMS

cost assignment - user should not be loaded with fixed costs not related to service - costs should be strictly proportional to use

feedback - user should be able to influence system by feeding back opinion and facts

indexing - options for professional searchers and for browsers should be provided - natural language and structured system aids such as the EXPAND function of DRB are desirable as also are stemsearching and hierarchical systems

ordering - ordering through the system should be easy

query language - should be available in versions to suit sophisticated and casual searchers - subset of natural langugage - in French and English

TECHNICAL PROBLEMS OF INTERFACING AUTONOMOUS SYSTEMS

The problems are in three areas: 1) computer-to-computer interfaces, 2) query language (and data base structure), and 3) indexing methods. Solutions to the first problem have long been sought at fundamental hardware levels using approaches such as that of ARPAnet based on interface processors. However another approach would be to treat the separate systems as black boxes whose outputs are known for specific inputs, as I have proposed as the metasystem concept (Amey, 1974). In this case, if the sign-on protocol and query language are known, the problem reduces to automatic interpretation of messages between systems which are not intrinsically incompatible. This approach also solves problem (2) and it is possible to make the system essentially transparent to the user, who merely decides which data base to use, while employing his own, local query language. At DRB, experiments are planned to test the metasystem approach, which is believed to have very wide application for the provision of retail online service routed from suppliers interested primarily in wholesale operation.

COMPATIBILITY OF INDEXING

Suggestions have been made that systems should employ universal numerical classifications which are, superficially, independent of language, and would appear to be particularly appropriate for computer manipulation. An elite of specialist users may indeed find this a satisfactory solution. However if one is to continue to seek the Holy Grail of direct accessibility to information by anyone at all, one needs at least the option to use natural language or authorized terms close to normal technical usage. Names of authors and of corporate sources are natural, universal search terms. To circumvent problems due

to mispelling or rigid formatting, either menu-display methods (EXPAND function of DRB RETRO system) or stem-searching is essential. Switching between thesauri has often been discussed (automatic replacement of user's terms by appropriate synonyms) but is expensive to implement on a universal scale. Natural language, machine-selected from the bibliographic data, should always be available as a last resort option in searching by unsophisticated users.

CONCLUSION

The system developed in DRB has already shown that we can already come close to meeting the criteria offered in this paper. Special problems of a defence system stand in the way of meeting some ideals. Non-defence suppliers have related problems due to copyright restrictions. These, however, are not impossible to overcome in computerized systems since measuring use of material and assigning royalty payments is then feasible. By using computer-communication utilities to retail services to individual users, all these costs can be equitably shared among those making use of the data.

REFERENCES

AMEY, Gerald X. "The Forgiving System...Human Aspects of System Planning." Proc. 35th Congress of FID Argentina, 1970

AMEY, Gerald X. "Information Networks and their Relation to National Policy." Proc. Int Conf. Inf Sci Tel-Aviv, Israel 1971

AMEY, Gerald X, SAMUELSON, Kjell and BORKO, Harold. Design and Management of Information Systems...Practical Guidelines. UNESCO Study (In Press) 1972

AMEY, Gerald X. "Use of a Metasystem to Interface System Users to many Autonomous Data Bases." Eight Annual Princeton Conf on Information Sciences and Systems, 1974

CURRIE, Jack D. "Micro-EDP System Solves Data Search and Security Problem." Canadian Data System pp. 40-42 Aug 1973

DEFENCE SCIENTIFIC INFORMATION SERVICE. DSIS Retrospective Search System User's Manual. DSIS Tech Note No. 12 1973

IRVINE, James J. "A Remote-Terminal Retrospective Search Facility Using a Hybrid of Microform and Computer Storage." Inf. Stor. Retr. Vol. 9 pp. 597-606 1973

KREVITT, Beth I and GRIFFITH, Belver C. Evaluation of Information Systems: a Bibliography 1967-72. Information part 2, Vol 2 No. 6, 1973

MEADOW, Charles. Man-Machine Communication. New York, Wiley-Interscience 1970

SCIENCE COUNCIL OF CANADA REPORT NO. 6. "A Policy for Scientific and Technical Information Dissemination." September 1969

SMITHERS, Peter R. The Current Status of Retrospective Searching in DSIS. DSIS Tech Note No. 14, 1974

SUMMIT, Roger K. "DIALOG: An Operational Online Reference Retrieval System." Proc. ACM 6th National Meeting pp. 51-55 1967

TABLE 1

SOCCRATES RETROSPECTIVE SEARCH SYSTEM OF DRB - RETRO

Main Parameters

Accounting	individual accouting for all users by host-system
Browsing	three levels - search-data (EXPAND function) online bibliographic data (microfilm cartridges) full text (microfiche)
Communications	provided by host-system - transparent to user
Cost assignment	DSIS responsible for file-formation user responsible for system use
Data base	search-data for 500,000 citations to defence technical literature
Feedback	comments and evaluation data can be recorded by user in the system
Host system	APL/PLUS computer-communications network of IP Sharp Associates Limited
Indexing	six search dictionaries available: Author COSATI code DSIS descriptors Keywords machine-selected from titles Source TEST thesaurus of the Engineering Joint Council
Ordering	optionally through system
Query langugage	French or English; command or conversational; normal Boolean combinations
Terminal	ASCII or IBM 2741-compatible

Fig. 4

```
      )LOAD RETRO
SAVED  10.14.02 03/22/74
DSIS RETROSPECTIVE SEARCH SYSTEM.... MK II
     (E)NGLISH -- (F)RANCAIS?
*** F                                                          Ex. 1
DESIREZ-VOUS VOIR LES NOUVELLES DU SYSTEME?
LES NOUVELLES SONT A DATE A PARTIR DU
11 FEVRIER 1974.
REPONDEZ (O)UI OU (N)ON
*** N
ENTREZ VOTRE CODE ET VOTRE MOT-DE-PASSE

***
VEUILLEZ ENTRER UN TITRE POUR VOTRE RECHERCHE
*** ESSAI
ECRIVEZ LE NOM DE LA PERSONNE POUR QUI LA RECHERCHE EST FAITE S'IL
N'EST PAS L'UTILISATEUR.
*** DAOUST AJ
COMMENCEMENT DE LA RECHERCHE
POUR LE MODE DE CONVERSATION ENTREZ 'CONV'
.01 AUTEUR BEAULIEU AJ
9 DOCS
.02 THEASURUS LASERS
548 DOCS
.03 (1 AND 2)
(1AND2)  3 DOCS
.04 AFFICHEZ 3
D1  3 (1)  *   (1AND2)
70-08045  R70/4023L
THE DEVELOPMENT OF THE TRANSVERSE EXCITATION ATMOSPHERIC
PRESSURE LASER
                1970

D2  3 (2)  *   (1AND2)
70-09248  R70/4624R
TRANSVERSELY EXCITED ATMOSPHERIC PRESSURE CO2 LASERS
AD-714 439      1970

D3  3 (3)  *   (1AND2)
72-06159  R72/3080L
REFER TO DSIS FOR DETAILS
                1972

.04 COUT                                                       Ex. 2
CONNECT    4.40 MINUTES    0.88 DOLLARS
CPU       10.93 SECONDS    3.27 DOLLARS
TOTAL                      4.16 DOLLARS
.04 FIN
     LA RECHERCHE EST TERMINEE.  VOS COMMENTAIRES SERONT APPRECIES.  Ex. 3
CES DERNIERS PEUVENT ETRE ENTRES APRES CE MESSAGE.  A LA FIN DE
VOS COMMENTAIRES OU SI VOUS N'EN AVEZ PAS TAPEZ 'FIN' SUIVI
D'UN RETOUR DU CHARRIOT.
PAS DE COMMENTAIRE
FIN
C'EST MAINTENANT TERMINE.  NOUS VOUS REMERCIONS POUR VOTRE
COOPERATION.  N'OUBLIEZ PAS DE REMPLIR LE JOURNAL.
TAPEZ ')OFF'
```