EDBASE 2 – Intelligent Keyboard System. Inclusive Access to Database Applications

Steve Abbott

Deputy Director, Essex Disabled People's Association, 90 Broomfield Road, Chelmsford, Essex, CM1 1SS

info@essexdpa.org

Within this paper an overview of the EDBASE 2 – Intelligent Keyboard system is presented. The system is an experimental universal data access application designed to be attached to any mainstream database table format. Interrogation of the database is through natural language, though the system can also interpret clipped sentences or jumbles of keywords. Support for disabilities is provided through the specific design of the interface, and, by application of a specially designed spelling correction system based upon unique auto-associative neural architecture that is capable of correcting the more complex errors that can result from a users impairment consistently interfering with their operation of the keyboard. The system is currently able to offer support to typing errors associated with a range of physical, sensory and cognitive disabilities. It is designed to operate on any standard type of Pentium computer and keyboard.

Keywords: Spelling Checkers, Neural Nets, Databases, Adaptive Interfaces, Disabled Access, Disability Support

1. INTRODUCTION - WHO ARE ESSEX DISABLED PEOPLE'S ASSOCIATION?

Essex Disabled People's Association (EDPA) is a County Wide charity first established in 1949. Its purpose has always been to provide services to disabled people in the County of Essex, but as the world has moved on, so EDPA has had to in order to continue to serve the needs of it's service users. It currently supports a network of 200 clubs providing social and support activities for disabled people; the total membership of these clubs is in excess of 9000 people. The EDPA Travel Voucher scheme now has over 600 members, all of whom receive £35 in vouchers per month to spend with over a hundred transport providers in Essex. These providers range from taxi companies to an airline. In addition EDPA has moved into education, providing one of the first formally recognised training courses in the country that allows disabled people to become access auditors. Once through the course they can provide a valuable consultancy service on a commercial basis to the many businesses seeking advice in the run up to the October 2004 Disability Discrimination Act deadline. At the core of EDPA however is its County Wide Disability Information Helpline. This valuable and much used service has been in operation for over 15 years, and is staffed by a team of 14 disabled volunteers. The Helpline is able to offer information and advice on the broadest range of disability issues, dealing with up to 20 cases per day on average.

2. CASE STUDY - THE DEVELOPMENT OF THE PROTOTYPE EDBASE SYSTEM.

Any information Helpline runs on the strength of the information that it holds. EDPA has an information resource in excess of 11000 individual records covering a broad range of topics. When physical space is limited, and storage of hard copy must take into account the reach of the volunteers using wheelchairs, it is obvious that the best information handling solution is going to be I.T. based. The EDPA Helpline operated four Microsoft Access databases that each covered aspects of the EDPA information resource. The difficulty however was that these systems were largely shunned by the volunteers in favour of the hardcopy versions. A major responsibility of the newly appointed Deputy Director was to ascertain why this should be happening, and, to quote the then Director of EDPA - "Get the volunteers to all use the computer".

In its simplest form this meant replacing the existing database systems with something that all the volunteers would be able to use, irrespective of their disabilities or level of computer literacy. Over a period of 3 months extensive consultation was carried out with the volunteers and a number of key factors were identified that seemed to be impeding the ability of the volunteers to use the existing I.T. systems. These were:

1. Small size of fonts used within the systems. The default setting in Windows is 10 point, which had not been altered.

2. Small size of buttons used, making them difficult to see and press with the mouse.

3. Leading on from the point above, having to use the mouse at all was an issue; many of the volunteers found this difficult, either due to symptoms of their disabilities, or relative inexperience with using the mouse.

4. The sheer quantity of controls, and the apparent complexity of it all. The use of hidden functions such as drop down lists or combo boxes seemed especially problematic.

5. The standard Windows colour scheme seemed especially inappropriate; black text on grey caused everyone some degree of difficulty. Admittedly it is possible to customise these settings; but when the level of I.T. literacy amongst your staff is fairly low, the volunteers will not have the skills to do so, or even realise that you could alter the screen format in the first place.

Taking the above list into account the following was created as a form of design brief:

1. Font sizes within the system to be no less than 16 point.

2. Colours used to be dark background with pale coloured text. This presents less information to the eye and as a result is more comfortable for everyone, not just those with sight loss.

3. Font type to be a clear non-serifed format.

4. No hidden or multifunctional controls anywhere in the system, and such controls as were used to be clearly marked as to their purpose - in English, not jargon. The controls should be of sufficient size to allow easy alignment with a mouse pointer, and with high colour contrast to allow easy identification of the controls from the background.

5. Use of keyboard controls, and short cut keys to remove the dependency on mouse altogether.

6. The simplest type of control identified for use in the system is a free entry text box; therefore some kind of support would be necessary to overcome the difficulties of typing errors.

A prototype system was built and installed on the EDPA Helpline in January 2001. All volunteers were able to effectively use the new system, and required little or no training to do so. The system's built in correction for spelling errors proved especially interesting and suggested a number of ways of improving the design of the prototype based around suspected patterns of keyboard misuse linked to impairment types. A system designed to take advantage of this theory could conceivably give greater flexibility of use, not to mention greater support to the volunteers using the system. These ideas were drafted as proposals, which were considered by EDPA's Director and Trustees. Formal research work commenced with academic support from Anglia Polytechnic University in July 2001.

Note – Case study is available in full from (Abbott, Davies & Mourtzoukous 2003)

3. THE EDBASE RESEARCH PROJECT.

The background of EDPA's Deputy Director was in the Earth Sciences, specifically in the field of data processing and remote sensing instrument calibration. The problem of correcting the misuse of the keyboard when being used for free text entry appeared to have distinct similarities to that of a remote sensing instrument requiring calibration. Specifically, to paraphrase between calibration theory and plain English, a distortion was occurring in the collected data (words being typed into the keyboard) caused by effects being generated by the way in which the data was being collected by the instrument. (The words being entered were not quite as they should be).

In order to achieve a working calibration on a remote sensing system, it is necessary to approach the problem with no preconceptions. Any solution that presents itself must be based entirely on what is in the data produced by that instrument. It is necessary to limit your view of the problem to the perspective of the instrument; anything else is irrelevant. The data in respect of someone typing at a computer can only be the keys that are being pressed in the order that they are used. The EDBASE prototype was upgraded to allow each key press to be recorded in a series of text files. This was carried out following the informed consent of the volunteers and staff at EDPA being obtained. The system was put in operation in July 2001. Since then every session that has taken place on the prototype database system has been recorded, typing mistakes and all. This data resource is over 4000 files in number, containing millions of keystrokes by the 14 volunteers over nearly 2 years.

The design of a second system can only be based upon what is contained within this data resource. The processing and analysis of this data has taken up the bulk of the time spent on the project since then; however, the results have been significant, and have allowed a second system to be designed. The new system is well under construction, and has been developed to an experimental state suitable for trial use.

4. THE NEW SYSTEM (EDBASE 2 - INTELLIGENT KEYBOARD)

Specific areas targeted for development in the research project included allowing a more flexible use of language in the entering of search parameters, and the ability to more effectively correct the typing mistakes resulting from free text entry being allowed in the system. The principle reason for this was the observation that these typing errors did not appear to be random, moreover, that the patterns that seemed to exist were linked to the type of disability of the user. The auto-corrective techniques employed in the prototype allowed the correction of approximately 70% of the mistakes being made. These mistakes varied from casual character substitutions to more complex phonetic errors common with dyslexia. This however was not viewed as being sufficient, the behaviour of the prototype would become erratic and confusing in the instances where it was unable to correct, often substituting a word it thought was acceptable; that in fact had no link whatsoever to the context of the original search term. Although this was certainly amusing to the volunteers most of the time, it was not ideal behaviour for a database application.

The analysis of the data collected showed that the typing errors did indeed not seem to follow an entirely random pattern. Based on this an attempt at creating a general classification system was made. This first attempt resulted in 32 classifications; that in turn allowed a series of rule based corrective systems to be designed and constructed. In tests these rule based modules proved incredibly successful with their intended mistake type. When the rule based system was pitted against common spellcheckers such as that employed in Microsoft word the difference in capability was enormous. For one mistake type using data collected from the EDPA volunteers to provide the sample errors Microsoft word was able to handle 29%, the targeted correction system achieved 100% and has continued to do so in every test since. However, the data equally suggested that the corrective modules on their own would not be enough; some type of management system would be necessary to run the corrective modules. Again the solution was sought from the data, using some of the more specialist data analysis and interpretative techniques employed in chemistry and geology, merged with methods employed in psycholinguistics and neural net / cognition theory a system was developed that would employ a type of neural architecture filled with the end results of a sophisticated data processing and analytical system to sculpt and form the data into a tool capable of analysing and making sensible suggestions for how any individual typing error may be corrected. This system has been set up in a skeletal form within the experimental version available today. In order to complete the neural data structure it will be necessary to obtain far more widely captured data. The current system has been based upon the records of 14 volunteers over a two year period, and although this represents a decent starting point, it can in no way be considered an accurate representation of the behaviours of the general population. For this reason the final year of this project will largely be taken up in trials and data capture operations. It is intended to capture such data from a variety of locations, including hospitals, colleges, libraries, and through the club networks of EDPA and equivalent organisations in other counties. The data will be fed into the processing system; which will in turn be used to fill the gaps in the management systems capabilities.

5. CONCLUSIONS.

The key to the entire EDBASE system is the data that has been collected, and the systems developed to process and handle that data. Everything that the EDBASE 2 Intelligent Keyboard is capable of doing is based upon its stored knowledge on the patterns within that data. No formula based statistical modelling techniques have been employed. The patterns once identified are used to construct targeted correction modules, which are managed by the neural net. The concept of a modular type correction system along these lines has been explored in outline in the study of (Seth & Kokar 2001), though that study did not identify the need for any type of management system controlling the modules directly, nor did it employ the data analysis techniques that form the critical core of the design and construction of the EDBASE 2 system. Neural nets as a means of providing innovative solutions to programming problems are not an especially new technique. Indeed (Hodge & Austin 2003) provides results of a spelling correction system employing a binary neural net approach, however the correction systems employed in that study rely on traditional algorithm based techniques including n-gram analysis, hamming distance calculations, and soundex encoding. What we have managed to do at EDPA; is to take a radically different approach, loading a hybridised type of neural net with the filtered essence of the typing behaviours of a number of individuals with a wide range of disabilities covering physical, sensory, and cognitive impairments. The resulting encoded neural data-structure can then be employed to correct typing errors within a free text entry database application. This, coupled with an innovative context sensitivity system allows for a highly flexible and intuitive means of accessing database systems. The techniques employed to develop this system, and the neural data-structure that manages it show signs of potential applications beyond the database system it is currently built into. The ability of the system to recognise spelling mistakes based on classifications can be potentially run in reverse as a diagnostic tool capable of spotting dyslexia, or other conditions, from a short keyboard typing session. In addition, the ability to identify patterns within an individual's typing behaviour equally suggests a potential security application based upon that behaviour being used as a continuously checked fingerprint, potentially far more reliable than a single password entry method.

The system has been designed as a software agent, fully capable of being installed over any standard database format accessible through the JET or DAO methods of access with a minimum of work. This makes the initial use of the system most comfortably aimed at Community Information Points and Library / Helpline facilities.

Trials of the EDBASE 2 system will be commencing shortly in 10 voluntary sector locations across the County of Essex. The experiences gained from these trials, added to the results of the planned intensive data capture activities, will allow the final completion of the system by the originally planned date of Autumn 2004.

REFERENCES

Abbott S., Davies C. and Mourtzoukos M. (2003). Usability for those with cognitive impairments: a case study and preliminary findings with dysphasic subjects. *Proceedings of HCI International 2003*, Volume 4 pp 817 – 821.

Hodge V. J. and Austin J. (2003). A comparison of standard spell checking algorithms and a novel binary neural approach. *IEEE Transactions on Knowledge and Data Engineering*, Volume 15 no 5 (September/October 2003).

Seth D. and Kokar M. M. (2001). SSCS: A Smart Spell Checker System Implementation using adaptive software architecture. *www1.coe.neu.edu/~kokar/publications/SpellChk.pdf*