

AMENDMENTS TO X00 NAME ENQUIRY TRANSACTION
FOR COMMUNITY CHARGE

1. Introduction.
2. Soundex Name Matching as it works now.
3. Initial Amendments.
4. Further Amendments to improve X00 performance.
5. Summary

Appendix A.

Appendix B.

Geoff Williams
Applications Support Team Leader
19.01.89

1. Introduction

The Name enquiry transaction X00 provides the ability to do imprecise searching of the database of names related to the Rates and Community Charge applications.

Once the Community Charge system is live the amount of effort X00 will be performing will increase approximately 5 fold, basically because there will be many more persons added to the database, for which any person can be responsible for more than one property and also liable for more than one account.

Because name enquiry transactions like X00 have to search through large numbers of similar names to identify the best match to the user's request, its response time has always been longer than most other transactions. As more names have been added over the years minor changes have been made to try to maintain response time. Now with a massive increase in prospect X00 processing requires to be redesigned to ensure the maintenance of acceptable responses.

Transactions H00 and C00 carry out similar functions on the names indexes related to Housing/GPA and the Creditors systems. Consideration will be given to applying the same design changes to these applications systems after the revised X00 has been implemented.

2. Soundex name matching as it works now

We currently have 3 names databases, two being linked to the same properties database. One is for Rates names only, another combines Housing/GPA names and the third holds Creditor names. The Creditor names database is independent of the properties database.

Each name on the database is assigned a unique 6 character key. Character 1 of the key contains a number in the range 0 to 9. In general if a name contains certain recognised buzzwords, such as LTD or SONS it is classified as an organisation name, and the first character of its key set to 0; otherwise it is classified as a person name, and assigned a key prefix of 1. This 0 or 1 value can be overridden for the purposes of security, and the name assigned a first key character in the range 2-9. Access to that name is then restricted to certain GPA programs and transactions. Names with a first key character greater than 1 are said to be members of the "Restricted Function Set".

Characters 2 - 4 of the key are the 'soundex code', generated from the first few 'significant' characters of the surname. The objective is to represent the "sound" of the name not its precise spelling, such that mis-pronounced or mis-heard names in telephone conversations can still be found. The method of generation is such that H.R. JOHNSON for instance will have the same soundex code as MR.J.K.JAMESON. So all people with the same soundex code as JOHNSON will be stored together, forming a 'soundex group'. A soundex group, then, is any group of names with the same first 4 key characters. Characters 5-6 contain a unique person reference number.

A name is submitted (a surname is the minimum requirement) for matching against the database.

The submitted name is passed to a standardisation routine, which calculates its soundex code and assigns 0 or 1 as a first character according to the contents of the name fields.

All names on the database that are within that soundex group are scanned, and matched against the submitted name.

If the submitted name was originally assigned a 0 key prefix, the corresponding soundex group with a 1 prefix is then scanned and matched; or vice-versa: so that the name is matched against both organisations and persons.

The best 99 matches are displayed.

Since the advent of the Housing application, the number of names within certain soundex groups have increased to alarming proportions. Whilst the majority of groups, 90% of organisation groups, and 60% of person groups, have a membership of less than 10, there are some groups which hold a membership of greater than 1000.

Large membership soundex groups are the main cause for a degradation of performance of the name matching transactions. Where submitted names belong to such groups, the transactions have to scan through to find the best possible match. It is also extremely likely that the users have cause to access these groups more highly than smaller soundex groups.

3. Initial Amendments

- 3.1 The Picking list will no longer display 'MULTIPLE ADDRESSES'. Where a street name has been submitted by the user and a match has been made on that address, then we want to display that address. Provided an address exists for a name, at least display the first address retrieved and if there proves to be more, display 'M/A' (highlighted) at the end of the address line. See Appendix A.
- 3.2 The number of Picking Lists will be extended to 99. A current restriction of the name matching transactions is that it can only display 11 pages (or 99 names), so that where a large soundex group is being displayed many names will not appear on the picking list. Currently, all 99 names are held in the SPA, so in order to allow up to 891 names, pages 12 to 99 will be held on the Paging database. (It is worth noting that there are currently 912 "M.. ALI"s on the Tower Hamlets Housing names database).
- 3.3 The Name Details display will be amended such that details of the selected name's account and property details will be shown. See Appendix B.

4. Further amendments to improve X00 performance.

i) Use of a secondary index to match on names with a known address.

For any person within the CC system there could be a large volume of address information (e.g. an in-borough address per account, perhaps a communications address for every account, an in-borough address for a property if the person is the responsible person, perhaps a communication address for a property if the person is the responsible person).

A great many searches of the names index are made with address details submitted. Currently X00 searches through a soundex group looking for a perfect match on the first four characters of the address details submitted, which is a laborious process when the soundex group is large (Although address details submitted should be street name only the users commonly supply a house number, which can throw name matching. X00 will be amended such that submission of street name only be highlighted to the user on the enquiry screen and that the street details field will only allow alphabetic characters).

Where street name details are submitted with name details and where soundex groups contain more than 99 people, a secondary index containing the soundex code and first four characters of the street name and the unique person reference number will be used to achieve quick access to any people that have the same address as the street name which has been submitted.

Where perfect matches are made, these persons will be displayed immediately (i.e. currently there are 9 names displayed on each page of the picking list. Using the sec. index if less than 9 names are matched then only those people will be displayed, otherwise the first page will be displayed with the option to retrieve the rest). When the only or last page of perfect matches have been rejected, the user will have the option to continue with partial matching. This will be communicated to the user by way of a message, e.g. "YOU HAVE SEEN ALL PRECISE MATCHES. PRESS ENTER IF YOU REQUIRE MORE PAGES OF PARTIAL MATCHES".

Subsequently, if the user requests more pages of matches, or there were no perfect matches, then a form of secondary matching will be done on the soundex group. This form of matching will also apply for searches of names in large soundex groups (> 99 names) when the user has not supplied any address details.

For soundex groups that contain up to 99 names the current method of name matching will apply.

ii) Use of a secondary index to perform secondary matching for large soundex groups

A secondary index comprising the soundex code and the first forename be used to perform secondary matching under some circumstances.

Where a first forename has been submitted by the user, the secondary index will be used to gain quick access to all persons in the submitted soundex group with the same first initial of their forename.

On occasions where the soundex group is large and the sec. index cannot be used, i.e. no first forename information was supplied by the user, then a message will be sent back to the user in the form "THIS IS A COMMON NAME. A FASTER RESPONSE TO MATCHING WILL BE ACHIEVED IF YOU COULD PROVIDE STREET NAME (ONLY) OR FIRST FORENAME (PREFERABLY BOTH). HIT ENTER IF YOU CANNOT PROVIDE MORE PRECISE MATCHING INFORMATION." On seeing this message, the user will have the opportunity to avoid name matching processing as it is performed now by providing the sort of information that will allow a sec. index to be utilised.

The following rules will apply:-

On submission of first forename and surname only, or on submission of prefix and first forename and surname only, or on submission of first forename and second forename and surname only, or on submission of prefix and first forename and second forename and surname only, the sec. index will be used to display up to 99 pages of names.

On submission of surname only, or on submission of prefix and second forename and surname only, or on submission of prefix and surname only, or on submission of second forename and surname only the user will have the option to utilise the sec. indexes by supplying street name and first forename details, otherwise name matching will be performed as of now, with up to 99 pages of names to be made available for display.

Where street name details have also been supplied then the above rules will apply after all perfect street matches have been found and the user requests more pages.

SUMMARY

The above amendments will greatly enhance name-matching transaction response times - a critical requirement of the Community Charge Working Group and will be of benefit to the other related Borough user departments.

The date of implementation for the Community Charge Name Enquiry transaction is end of June.

Picking-List Display - format

10	20	30	40	50	60	70	80
RECORDS ARE HELD FOR							
1 ←	(name)		→		PAGE MM OF NN		
5 ←	(address)		→				
2 ←	→		→				
3 ←	→		→				
4 ←	→		→		M/A		
10 ←	→		→		M/A		
5 ←	→		→				
NO ADDRESS EXISTS.							
6 ←	→		→				
15 ←	→		→				
7 ←	→		→				
THIS IS AN ALIAS NAME. SEE DETAILS.							
8 ←	→		→				
THERE ARE ALIASES OF THIS NAME. SEE DETAILS.							
20 ←	→		→				
OPTIONS ←							
25 ←	→		→				

Name-Details Display - Format

10	IS THIS THE PERSON/ORGANISATION REQUIRED?	50	60	70	80
5	NAME: <				
	ALIAS: <				
	PRIME: <				
	CHARGE A/C NNNNNNN	CC PAYER FOR: <			
10	CHARGE A/C NNNNNNN COMMUNICATIONS TO: <				
	PROPERTY REF. NNNNNNN	RESPONSIBLE FOR: <			
15	PROPERTY REF. NNNNNNN COMMUNICATIONS TO: <				
20					
	OPTIONS <				
25					