NU Science Journal 2006; 2(2): 139 - 150

Hybrid Name Matching Methods for Rule Based Thai Naming System

Chakkrit Snae* and Michael Brueckner

Department of Computer Science, Faculty of Science, Naresuan University, Phitsanulok 65000 Thailand
*Corresponding author. E-mail: chakkrits@nu.ac.th

ABSTRACT

Names are important in many societies, even in technologically oriented ones which use ID systems or other ways to identify individual people. There are many elements of personal names which vary in and between the different cultures. Names such as personal surnames are the most important as they are used in many processes, such as identifying of people, record linkage and for genealogical research as well. On the other hand, variation of names can be a major problem for the identification and search for people, e.g. web search or security reasons. We show name variations for different cultures to guide the implementation of a rule based naming system, currently worked out for Thai names. We discuss characteristics of the LIG (Levenshtein, Index of Similarity Group (called ISG), and Guth) algorithms which help to find reasonable variants of names. A further benefit of this process would be an optimization for name searching.

Keywords: personal names, name variations, name matching, rule based system

INTRODUCTION

Names are used for identifying persons, places, things and even ideas or concepts. Names serve for labelling of categories or classes and for individual items. They are properties of individuals which are of greater importance in most communities. In technological oriented societies such as modern Western the reference between names as a label and the person is not as obvious as in small tribal societies. This is especially true where names are stored within large information systems. This includes government, medical, educational and even commercial records which are kept about individuals. Names are the most important referrer to a person even if there are numbering systems like ID numbers because such systems are not universal. Names are often queried in a different way than they were entered. Personal names lead to many problems with regard to data retrieval because names are also subject to multiple variations not only between different cultures and writing system but in a specific culture as well. Names represent complex lexical structures which have to be handled systematically for data entry, storage and retrieval in order to get sufficient recall or precision the retrieval process in.

In this paper we present a first account of our findings on elements of personal names in different cultures with their respective conventions and variations, rules for constructing Thai names, and algorithms of name matching to overcome name variations. Here we present a hybrid name matching procedure which is based on

probabilistic phonetic and sound variation recognition with the help of an expert system and which can deal with multicultural names as well. This procedure is used in the Thai naming system.

This paper is organized as follows: First we give a description of names and their variation. After that we present some basic characteristics of the name matching algorithms of our choice, the LIG algorithms, as well as the rule based system for naming applied to Thai names. The last section shows the conclusions of our study and further work which has to be performed.

Background

1 Meaning of Names

Names for individuals are often called proper names, for humans sometimes also anthroponyms. Names for places are called toponyms, for bodies of water hydronyms, for ethnic groups ethnonyms, for metaphors metonyms and so on. Names are more than just strings of characters. Names show important information such as titles, gender, marital status, and even birthplace. For this names provide different elements, which may differ between cultures. They also undergo variations such as phonetic and alternate spellings. This problem can be made clear in this way: how do we know that differently spelled or pronounced names belong to the same person?

Each culture has a set of conventions which govern the appearance and function of names as well as a range of permitted variations in its naming system, e.g. surname inheritance. Some cultures show the marital status very clearly, others do not refer to it, cf. Thai and German differences in stating the marital status for female. The categories of name variation are generally the same across all cultures, the difference is in the realization of the variation.

1.1 Naming for Identity and Security

From the technical point of view we want to link and match as many names as possible with the correct individuals. If we deal with individuals of the same name, e.g. John Smith, we have to establish a second identifier at least. This can be – and is in many cases – a temporal element, like the date of birth, which is an individual and unchanging property of the person. Another way to circumvent the problem is to establish numbering systems, like ID numbers. Systems of numbers or other ciphers can be generated within individual organisations. It is not likely that the resulting ID numbers will be the same in different organisations. The numbering may have limitations as well, e.g. the individual health care setting (e.g. within a hospital or district) or, in principle, more widely (e.g. the National Health Service number). In the past, the National Health Service number in England and Wales had serious limitations as a matching variable, and it was not widely used on health-care records. With the allocation of the new ten-digit number throughout the NHS all this has been changed (Gill, 1997).

Although numbering systems are simple to implement they can lead to different errors in recording, transcripting, and keying. So we have to take into account methods which reduce these errors and facilitate good quality of data entry and retrieval. One such method uses a checking device such as check-digits (Wild,

1968; Hamming, 1986). When we are not able to use unique numbers or ciphers, natural matching variables are the person's name, date of birth, sex and perhaps other supplementary variables such as the address with postal code and place of birth, which are used in combination for matching. Recently, it has been suggested that this simple code could be extended for security critical places (e.g. airports, checkpoints etc.) with biometric marker information extracted from person identifier information e.g. fingerprints/iridograms.

1.2 Elements of Personal Names

The following Figure shows typical elements of personal names together with potential variations and sources of choices, e.g. dictionary of given names.

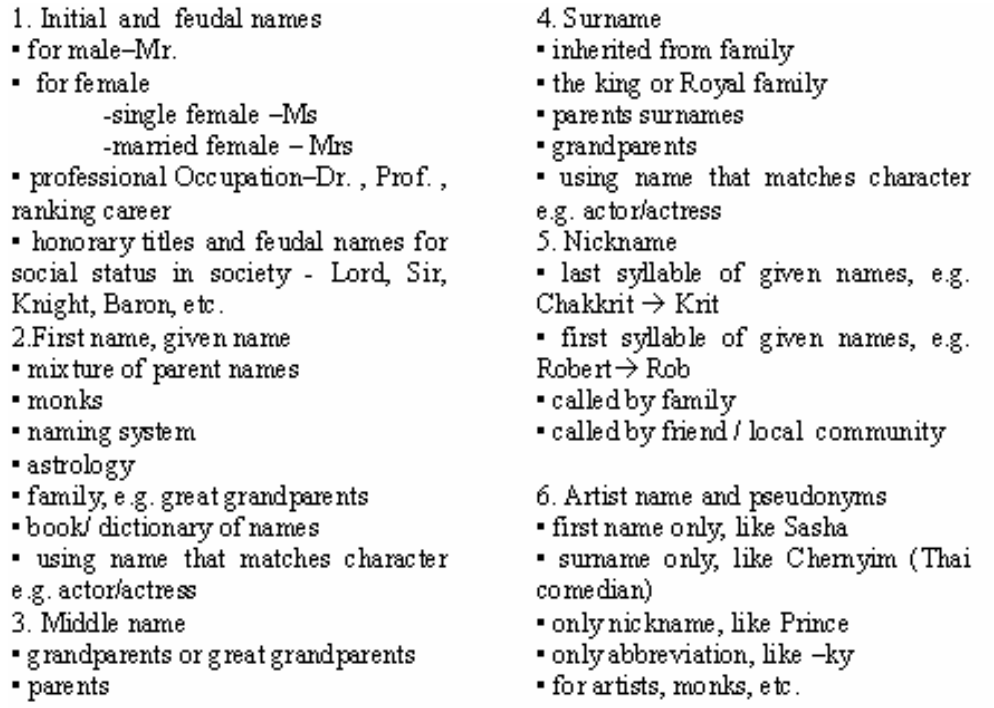


Figure 1 The elements of names

1.3 Variation of Names

Name variation is one of the major problems in identifying people, because it is not easy to determine whether a name variation is a different spelling of the same name or a name for a different person. Most of these variations can be mainly categorised as follows (Bouchard and Pouyez, 1980; Branting, 2001; Dematteis *et al*., 1998).

1.3.1 Spelling Variations

Spelling variations rely on the assumption that the source and target names are strings which differ because of errors or transcription differences (e.g. different pronunciation). Spelling error patterns can be taken into consideration and single-error misspellings (mistyping) can be categorized as follows (Jurafsky and Martin, 2000): (1) insertion, e.g. BROWN as BROWMN, and MCMANUS as MACMANUS; (2) deletion or omission, e.g. BROWN as BOWN, and ROBBIN as ROBIN; (3) substitution, e.g. BROWN as BTOWN, and SMYTH as SMITH; (4) transposition, e.g. BROWN as BRONW, and BREADLEY and BRAEDLEY. Generally such variations do not affect the phonetic structure of the name but still cause problems in matching names. These variations mainly arise from misreading or mishearing, by either a human or an automated device. These can include interchanged or misplaced letters due to vowel replacement (EVANTUREL as EVENTUREL), consonant replacement (LEBRE as LETRE), consonant doubling (MAUFET as MAUFFET), very different spellings (LEWIS as LOUIS), and problematic transcription (GARWOOD as YARWOOD) (Winchester, 1973).

1.3.2 Phonetic Variations

Phonetic variations depend on the dialect or pronunciation conventions of the speaker. For example, the nickname Pooh, as it is spelled in English, would be spelled in German as Puh. Where the phonemes of the name are modified, e.g. through mishearing, the structure of the name may be substantially altered. MAXIME and MAXIMIEN are related names but their phonetic structure is very different. Indeed, phonetic variations in first names can be very large as illustrated by ADELINE and its shortened form LINE.

1.3.3 Character Variation

The problem created by capitalization, punctuation, spacing, qualifiers and abbreviations (Branting, 2002) can be shown as follows:

Capitalisation, e.g. brown and Brown; SMITH and Smith
Punctuation, e.g. WILL SMITH and WILL-SMITH; SMIT and S.M.I.T
Spacing, e.g. YOUNGSMITH and YOUNG SMITH
Qualifiers, e.g. WILL SMITH and WILL SMITH YOUNG
Abbreviations, e.g. ROB and ROBBIN; BOB and BOBBY

1.4 Culture of Names / Naming Conventions and Customs in Different Cultures

The issue of name variations becomes more problematic when dealing with names from other cultures because the sorts of variation that are permitted may not be the same as those permitted in English. Names vary between cultures which for a long time has been an obstacle for creating a single method for automatic name processing.

For example, within each of the following cultures (Korean, Arabic, Hispanic, and Hungarian) all the names given are permitted variants of the same name except

the last one (Dematteis *et al*., 1998). How to find out which of the examples below are given names and surnames?

PARK DOE REE / PAG TO NI / TO NI PAG (Korean), MOHAMMAD ALI ABD EL NADIR NUR EL DIN / IMHEMED ABDUNADEER NOOREDDINE / MHMD NUR ABD AL NADER (Arabic), ENRIQUE CESAR VELEZ ARGUETA ENRIQUE BELES, QUIQUE VELEZ A. E. C. ARGUETA (Hispanic), Eoetvoes Lorant / Roland Eoetvoes / Eoetvoes Roland (Hungarian)

In each case, the final name would be considered an unacceptable variation of the name under consideration; it would be another name (Dematteis *et al*., 1998).

1.4.1 Order of Names

Name order is crucial; although the spelling variants of the name elements in the final name are acceptable, their order is not. In Eastern naming system, e.g. Korean, Japanese, and Chinese, the family name appears in the leftmost position and cannot move to the rightmost position. This applies to Hungarian personal names as well. In Western system, e.g. English, German, French, Spanish, Italian, and American, the family name appears in the rightmost position.

For example, in Hispanic names, the family name is the next to the last element (VELEZ); the rightmost name (ARGUETA) may be dropped, but not the family name. ARGUETA would therefore refer to another family, if it occurs alone (Dematteis *et al*., 1998).

As the function of personal names is not only to distinguish between individuals but also to serve as a help for indexing we have to mention the different conventions for ordering names. In Western order we change the position of given names and family names for indexing. In Eastern we do not have to change because the order for family names is already there. Thai ordering of names is distinct to these because Thai names are sorted according to the first names. There is no need to change the order of first name and family name as well.

1.4.2 Transcription of Names

English spelling with its many-to-many sound/letter correspondences contributes to the problem of Romanization of non Western names. Dialectal differences, historical and phonetic spelling all make the English names somewhat unpredictable. The latter is even the case for English names. The following examples are given by Reaney and Wilson (1997).

COLWELL, COLWILL, COLLWELL
LEA, LEE, LEGH, LEIGH, LEY, LEYS, LAY, LAYE, LYE
THOMPSON, THOMSON, TOMSEN, TOMSON
WORCESTER, WORSTER, WOOSTER, WOSTEAR

Spelling variations are especially prominent in names from non-Roman writing cultures when such names have been transcribed to Roman characters, e.g. the Romanised form of an Arabic name: NOOR EL DIN, NURELDIN, and NUREDDINE.

For the Thai writing system which is very complicated there exists an official standard called Royal Thai General System of Transcription (Wikipedia 2005) which is used for rendering Thai names into the Roman alphabet. It uses only straight letters for vowels, diphthongs and aspirated consonants. It does not indicate the length of a vowel and the five different tones. From Chulalongkorn University, Bangkok, there is also an automated tool available which transcripts Thai names or terms into Roman letters. It is called "Thai Romanization" (Aroonmanakun 2004).

In many cultures, available standard transcription systems are not used or are used inconsistently. The range of variation found in distinct instances of the same name is therefore not fully predictable from such systems. In Arabic, for example, although there are transcription systems used by libraries and other official agencies, transcription tends to be far less predictable and highly inconsistent, even with a single individual. For example, an individual whose name is "ABD EL NADIR" may romanise the name on one occasion as ABDUL NADEER and on another as ABDUNNADIR. Both name representations are "correct" and can be said to be accurate Romanisation of the same Arabic name. Even in cultures in which transcription systems provide a reliable standard, personal interpretation, accommodation to the spelling of another culture or perceptual confusion can cause the spelling to deviate from the standard. Thus, for example, the Thai name GOFF will vary with KOFF, because G and K are Romanisation alternatives from different transcriptions systems. An observed variant of GOUGH, however, are GOFF and GOFFE, representing the influence spelling of English surnames by Reaney and Wilson (1997). The Thai Romanization tool gives KOP; on the other hand KOP is a correct transliteration of two different Thai names: กอฟ and กอล์ฟ.

METHODOLOGY

2 Implementation of Hybrid Algorithms and System Concept

Many techniques have been used to cope with the important problem of matching variant names. However, most of these techniques were developed for general word matching and as a result they are not optimised for personal names matching. Spelling as well as phonetic variations combined with cultural aspects are the more challenging problems for automated multicultural name matching systems. Most systems today use specific techniques for name matching rules and specific variations (e.g. Guth (1976) and Levenshtein (1965) algorithms are spelling/string analysis based), whereas Soundex (Winchester, 1970) and Phonex (Lait and Randell, 1998) are phonetic/sound based algorithms. However, most researchers have tried to implement a method which can deal with the culture of names and naming system which are meant to overcome the ethnic problems, e.g. NameX (NameX), Varispell (Varispell).

2.1 Matching Algorithms for Names

The difficulty of the name matching task and the requirements for an effective algorithm to perform this task, both depend on the type and degree of name variations which occur. More recently published name matching techniques are

either of the composite or hybrid form (Snae and Diaz, 2002) and several novel hybrid algorithms (e.g. LIG2, and LIG3) have been developed for specific purposes. All the name matching algorithms encountered in the literature and presented in this paper are based on alphabetic and/or phonetic similarity and/or name transformations (e.g. forename abbreviations) but may use a variety of distance and other metrics for representing the match. From an initial search of the literature, Snae (Snae and Diaz, 2002) distinguished four types of algorithms and implemented them using the C programming language:

1. spelling/string analysis based algorithms (e.g. Guth and Levenshtein),
2. phonetic/sound based algorithms (e.g. Soundex, Metaphone, NYSIIS, and Phonex),
3. composite methods (spelling or sound, e.g. SIMPLEX, and ISG),
4. hybrid approaches (spelling and sound, e.g. LIG algorithms).

A hybrid algorithm combines phonetic and spelling based approaches using similarity measure as probability is called LIG algorithms (e.g. LIG3) (Snae and Diaz, 2002). The LIG algorithms have the best performance in term of producing most accurate true matches, overcoming name variations, and increasing the hit rate. They have proved to be more accurate than other methods in the literature (Snae and Diaz, 2002). The advantageous characteristics of these algorithms can be summarized as follows: (1) simple design, which can lead to accuracy improvements without decreasing the performance, (2) use of probabilistic similarity measures based on distance and weight, (3) increase correct positive and reduce negative matches to maximise the overall accuracy, (4) provide phonetic tuning to address multi-cultural names without depending on the language.

For these reasons we use LIG algorithms for name variations and matching in the proposed naming systems, e.g. Thai naming system. For that we use a dictionary database of more than 8.000 Thai names which contains not only the spelling, but also the meaning and correct pronunciation. In situations where names follow the rules but do not have a meaning we compare the name with similar names in a dictionary database and check for similarity above a specific threshold. Then the user can select the best name from the resulting list of names.

2.2 Rule Based Naming System for Thai Names

In this section we will show the three methodologies for naming as well as the basic rules for Thai names.

2.2.1 Naming Methodology

The way of naming can vary, e.g. naming by monks, grandparents. Since former times names are very important to people. Naming from the past to the present has been continuously developed and has developed a variety of patterns. Each pattern has it own rules depending on local places and the belief that has been developed until the present. The basic goal of naming is to provide a good fortune and progress during life. Most first names have a meaning. Three methodologies are briefly described in the following.

- Principal naming using Thai astrology is widely used since the past. Because it uses birth day to form the name. This is a belief that the individual has a set of 8 attributes called name of the angles referred to in Thai astrology. These attributes influent each person's livelihood, fortune, etc. The attributes refer to Servant >Age> Power> Honour> Property> Diligence> Patron> Misfortune. Each attribute has it own letters which can be used for constructing names.
- Principal naming using numeric methodology: Each letter has distinct numbers which can be added and have according values. These values represent low or high characteristics. The method can always be used along with naming both first names and surnames by increasing the value of first name and surname. Thus using numeric methodology can be used to increase "power" in names and check for better names.
- Principal naming which uses the traditional calendar is considered by Thai fortune tellers as the best method of anticipating the horoscope or destiny of people. This methodology takes day, month and year of birth including the time of birth to calculate the personality according to astrology. The results of this prediction are defined to tell the fate and personality thoroughly in the future.

2.2.2 Basic Rules of Forming Syllables

Syllables are aggregated to names which sound good or aimed at good fortune according to the three methodologies mentioned above. As a consonant can not stand alone in Thai language and personal names we consider rules for vowels only. The order is:

Vowels can come first or can be followed by a first consonant, e.g. Ek
Vowels can follow a first consonant without a final consonant, e.g. Ka
Vowels that can not have final consonant, e.g. Tam, Tua
Vowels that need final consonant, e.g. Kak

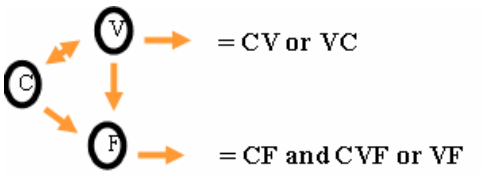


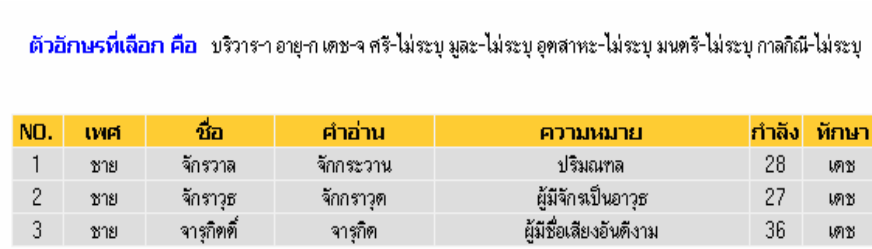
Figure 2 Forming of Thai syllables

Example of construction of Thai syllables using Thai Romanization 1.10 unicode (CU 2004) according to Figure 2: กา (Ka) = CV, เอก (Ek) = VC, กก (Kok) = CF, กัก (Kak) = CVF, เอก (Ek) = VF.

RESULTS AND DISCUSSION

We have constructed and implemented a web-based Thai naming system which offers two basic ways to come to "good" Thai names according to the first methodology mentioned above. The system gives us the letters for each date of birth. Figure 3 shows the letters for Sunday and then user chooses the vowel "า" from attribute Servant (บริวาร), letter "ก" from attribute Age (อายุ) and letter "จ" from attribute Power (เดช). We use these letters to construct names based on the basic rules (see Figure 2). Then we compare the different results with our database and return the hits. The user will be able to choose from a list of resulting possible names according to their respective meaning, e.g. จักรวาล, จักราวุธ, จารุกิตติ์ (Figure 4).

Figure 3 Thai naming system returns the letters for the day of birth which is Sunday (screenshot)



ตัวอักษรที่เลือก คือ บริวาร-า อายุ-ก เดช-จ ศรี-ไม่ระบุ มูละ-ไม่ระบุ อุตสาหะ-ไม่ระบุ มนตรี-ไม่ระบุ กาลกิณี-ไม่ระบุ

NO.	เพศ	ชื่อ	คำอ่าน	ความหมาย	กำลัง	ทักษา
1	ชาย	จักรวาล	จักกระวาน	ปริมณฑล	28	เดช
2	ชาย	จักราวุธ	จักกราวุด	ผู้มีจักรเป็นอาวุธ	27	เดช
3	ชาย	จารุกิตติ์	จารุกิต	ผู้มีชื่อเสียงอันดีงาม	36	เดช

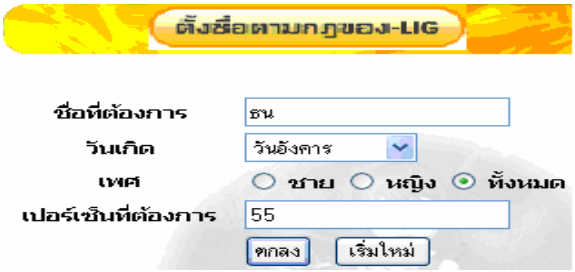
Figure 4 Thai naming system return results from constructing names based on the letters for Sunday and male gender using basic rules of syllables (screenshot)

The LIG algorithms help to combine different letters, syllables and names (e.g. parents' names) into a new name with a meaning. For example, users would like to include the letters th and n (ธ and น) or thn (ธน) from parents' name into baby names, so they have to select the day of birth (e.g. Tuesday) and specify the letters to be included in the new name (see Figure 5). Figure 6 shows the results from using the LIG 3 algorithm to match chosen letters and names in the database and to calculate the percentage of names that contain the selected letters.

The LIG 3 used in this paper is described by Snae (Snae and Diaz, 2002):

$$\text{LIG 3} = \frac{2I}{2I + C}$$

where I is the number of identical letters in the two names calculated using the Levenshtein method; and C is a Levenshtein cost calculation (Levenshtein, 1965). For example, the given name (1) ธนัท (Figure 6) compared to the chosen letter ธน from Figure 5 and the calculation of LIG3 is 2(2)/2(2)+2 = 4/6 = 0.66 *100 = 66%. C is 2 because we have to insert two letters ท์.



55% threshold for LIG 3

Figure 5 Thai naming system using LIG 3 algorithm

แสดงรายการค้นหาวัน **อังคาร** เพศ **ทั้งหมด**

เลขที่	เพศ	ชื่อ	อ่านว่า	ความหมาย	พักษา	วัน	%
1	ชาย	ธนัท	ทะนัด	ผู้ให้ทรัพย์, ผู้ร่ำรวย	เดช	อังคาร	66
2	หญิง	ธนาภา	ทะนาพา	รุ่งเรืองด้วยทรัพย์	เดช	อังคาร	57
3	ชาย	ธนิสร	ทะนิสอน	เจ้าแห่งทรัพย์	เดช	อังคาร	57
4	ชาย	ธเนษฐ	ทะเนด	ร่ำรวยสุด	เดช	อังคาร	57
5	หญิง	ธยานี	ทะยานี	ผู้มีปัญญาแห่งพินิจ	เดช	อังคาร	57
6	หญิง	ธัญมน	ทันยะมน	มีจิตใจดียิ่ง	เดช	อังคาร	57
7	หญิง	ธันยา	ทันยา	มีบุญ, มีโชค	เดช	อังคาร	57
8	ชาย	ธาวิน	ทาวิน	ผู้บริสุทธิ์, วิ่ง, ว่องไว	เดช	อังคาร	57
9	ชาย	ธฎิน	ทะติน	ชื่อพระศิวะ	เดช	อังคาร	66
10	ชาย	ธนชิต	ทะนะชิด	ชนะด้วยทรัพย์	เดช	อังคาร	57
11	ชาย	ธนตล	ทะนะตน	บันดาลทรัพย์	เดช	อังคาร	66
12	ชาย	ธนตุล	ทะนะตุน	มีทรัพย์นับไม่ได้, มีทรัพย์มาก	เดช	อังคาร	57
13	ชาย	ธนทัต	ทะนะทัด	มีทรัพย์	เดช	อังคาร	57
14	ชาย	ธนวัต	ทะนะวัด	มีทรัพย์	เดช	อังคาร	57
15	ชาย	ธนัช	ทะนัด	เกิดจากทรัพย์, คนร่ำรวย	เดช	อังคาร	66
16	ชาย	โยธิน	โยทิน	นักรบ, ผู้ชนะ	มนตรี	อังคาร	57
17	ชาย	วิธาน	วิทาน	กฎเกณฑ์,ระเบียบ	มูละ	อังคาร	57
18	หญิง	เวธนี	เวทะนี	สว่าน, มีปัญญาเฉียบแหลม	มนตรี	อังคาร	57

Figure 6 Results of Thai naming system using LIG 3 algorithm (screenshot)

CONCLUSION

We have used Thai astrology as a naming methodology, the LIG 3 algorithm for personal name matching and the basic rules for forming syllables in Thai to construct the rule based naming system. Our proposed system uses the hybrid name matching algorithms to return the variants of names from a database with the relative probability of their similarity. The advantage of this process is to improve searching algorithms for multicultural names in databases as well as in the internet.

Currently we are designing a system for multicultural name matching called NARESUAN-M^2 (Name Recognition Expert System Using Automated Name Matching Methods).

Our future intention is to use ontologies to check the different principal names according to Thai astrology by implementing an indexed database system of names from Thai dictionary for every day of a week. Furthermore we want to combine this method with the numeric methodology mentioned in this paper. A primary objective here would be to study how ontologies and algorithms can help in deciding which rules of naming system have to be implemented. This will also require an investigation into how ontologies covering the different elements of names can be merged.

REFERENCES

Aroonmanakun, W. (2004). Thai Romanization, Retrived November, 19, 2005, from http://www.arts.chula.ac.th/%7Eling/tts/.

Bouchard, G. and Pouyez, C. (1980). Name Variations And Computerised Record Linkage. *Historical Methods*, *13*(2), 119-125.

Branting, L.K. (2002). Name-Matching Algorithms for Legal Case-Management Systems. Refereed article in: *The Journal of Information, Law and Technology (JILT)*.

Dematteis, K., Lutz, R. and McCallum-Bayliss, H. (1998). Whose Name Is It: Names, Ownership and Databases. Originally written for: 1998 Annual Meeting American Name Society San Francisco, CA.

Gill, L. E. (1997). OX-LINK: The Oxford Medical Record Linkage System, Complex linkage made easy, Record Linkage Techniques. In: *Proceedings of an International Workshop and Exposition*, 15-33.

Guth, G.J.A. (1976). Surname Spellings and Computerized Record Linkage. *Historical Methods Newsletter*, *10*(1), 10-19.

Hamming, R.W. (1986). Coding and Information Theory (2nd ed.) Englewood Cliffs, NJ: Prentice Hall.

Jurafsky, D. and Martin, J.H. (2000). *Speech and Language Processing*, Prentice Hall.

Lait, A.J. and Randell, B. (1998). An Assessment of Name Matching Algorithm. *Society of Indexers Genealogical Group, Newsletter Contents, SIGGNL issues 17*.

Levenshtein, V. I. (1965). Binary codes capable of correcting deletions, insertions and reversals. Doklady Akademii Nauk SSSR, *163*, 845-848 (trans. Soviet Physics Doklady 10, 707-710).

NameX, Retrived November, 20, 2005, from http://www.imagepartners.co.uk/Thesaurus/AboutNameX.htm.

Reaney, P.H. and Wilson, R.M. (1997). *A Dictionary of English Surnames*, Oxford: OUP.

Snae, C. and Diaz, B.M. (2002). An Interface for Mining Genealogical Nominal Data Using the Concept of linkage and a Hybrid Name Matching Algorithm, *Journal of 3D-Forum Society*, *16*(1), 142-147.

Varispell. Alphabic Multi-Lingual Search & Recognition, Retrived November, 23, 2005, from http://www.alphabic.com.

Wikipedia. (2005). Royal Thai General System of Transcription, Retrived November, 19, 2005 from http://en.wikipedia.org/wiki/Royal_Thai_General_System_of_Transcription.

Wild, W.G. (1968). The Theory of Modulus *N* Check Digit Systems. *The Computer Bulletin*, *12*, 308-311.

Winchester, I. (1973). On referring to ordinary historical persons. In: *Identifying People in the Past*. E.A. Wrigley (Ed.), 17-40.

Winchester, I. (1970). The Linkage of Historical Records by Man and Computer: Techniques and Problems. *Journal of Interdisciplinary History*, *1*, 107-124.