Thai and English Electronic Dictionary for Android Mobile Phones Using Python Language

Chakkrit Saengkaew

Abstract—The world today can communicate in a wide dimension. There is a connection among groups of people of different races. One obstacle of the communication is the problem about language. Even though there are many courses in English language in Thailand, there are still a lot of Thai people who cannot communicate in English. The developer recognized this importance and decided to develop a tool which can help people (Thais and foreigners) connect. Android Mobile phones are widely-used devices and they have a lot of potentials in becoming an electronic dictionary (both Thai-English and English-Thai) so that android mobile phone users can look for words and translations. This way, Thais and foreigners can study, look up, understand, widen opportunities and reduce the gaps during communication. Programs and databases on android mobile phones do not need to be accessed online because users can look up Thai and English words on their own android mobile phones.

Index Terms—Android mobile phone, Android operating system, electronic dictionary, mobile application, NECTEC Lexitron, Python Language, SQLite DBMS, thai language, SL4A.

I. INTRODUCTION

English is considered the universal language for communication between people in the world who use different languages. In Thailand, there is a communication between Thais and foreigners all the time. During the moment when the word is not recognized, we tend to use a dictionary [1] to look up the definitions or the foreign words. It takes time to look up the words online [2] because there is a restriction about the computer size and the internet connectivity. There are electronic dictionaries available but they are quite expensive. The developer decided to build a dictionary on an android mobile phone because android mobile phones are devices widely used. The dictionary which the researcher developed can translate Thai into English and English into Thai. Thai is a unique language in that there are 44 consonant letter and 32 vowels along with 4 tone marks.

Some consonants can work as final consonants and vowels. Some vowel forms require other vowel forms so that they become words. Therefore, dictionaries on android mobile phones are important for Thais and foreigners during communication.

This research aimed at designing and building dictionaries on android mobile phones. The program has 2 main components that are database for vocabulary and program to look up words. The database consisted of vocabulary, translations.

The data were stored in the form of Table (SQLite DBMS) and interface using python language on SL4A layer running on android mobile phones. The program to look up words interacts with users and is used to search for words in terms of SQL on android mobile phone display. This journal will explain the procedures and summarizes in the last section.

II. LITERATURE REVIEW

Yuen Poovarawan [3] did research on analysis of Thai words and built a dictionary by choosing randomly word and sentence examples extracted from books, newspaper, journals, magazines, letters, official letters and general reading books, excluding books about literature and academic texts translated from other languages. Therefore, words from foreign countries are not included in LEXiTRON [4], an online dictionary developed by Nation Electronics and Computer Technology Center or NECTEC, and Nation Science and Technology Development Agency or NSTDA. This dictionary contains frequently used words in documents. The current database consists of 79,000 English-Thai translations and 51,000 Thai-English translations [5]. However, LEXiTRON is limited to those who use can access the internet, resulting in limitations for android mobile phone users who cannot access the internet. This project was initiated to make LEXiTRON available for android mobile phones. At the same time, the accuracy of the data is still the same as the database. The contents are the same. Only the format is changed. The overall picture of the research is shown in Figure 1.

![Fig. 1. Shows the overall idea of Thai and English electronic dictionary for android mobile phones](image-url)

Today’s smartphones support an abundance of different programming options. The researcher focused our work on the Android platform, as it is still the most prominent platform by far [6], with the largest number of available devices.

However, despite today’s abundance of feature-rich mobile phone hardware and powerful software platforms, creating applications that leverage the platforms’ potential is...
still a time consuming process that challenges non-expert developers by requiring in-depth know-how. [7]

With the rise of the iPhone and the Android platform released by Google, the share of smart phones has been constantly growing[8]. A key advantage of smart phones is the ability to access the Internet while on the go [9]. Interaction wise however, smart phones still fight a number of teething problems, including the cumbersome entry of data.

While a number of scripting languages are available for mobile phones, there is usually a drawback involved. Either they are in early development (Python[10], Perl[11], JRuby[12], Lua[13],php[14]).

III. CONVERTING VOCABULARY DATABASE

Converting Thai and English vocabulary database required programming and working on microcomputers via Python. The database of LEXiTRON [4] which is called LEXiTRON DB was distributed in the form of XML and then converted to SQLite DBMS format. The developers used Python Language running on SL4A Layer of the android environment [10] because Python is stable and has a lot of libraries. It is also platform-independent and freeware. The procedures are shown in Figure 2.

Fig. 2. Shows the procedure of converting database on microcomputer

The vocabulary database was converted from XML format which is specific like

```xml
<Doc><eResearch>yak</eResearch><eEntry>yak1</eEntry><tEntry>วัวป่าขนยาวในทิเบต</tEntry><ecat>N</ecat><id>82946</id></Doc>
```

The full source code for convert English Thai XML Format to SQLite DBMS which is specific like

```python
>>> t = open("/home/dictionary/etlex","r").read().split("</Doc>
<Doc>")
```
Fig. 3. Shows the procedure of converting XML DB to SQLite DBMS

Sample source code for Insert English Thai XML Format to SQLite DBMS which is specific like

```python
# -*- encoding: utf-8 -*-
import sqlite3

DB_NAME = '/mnt/sdcard/etlex.db'

database = sqlite3.connect(DB_NAME) # Create a database file
cursor = database.cursor() # Create a cursor

cursor.execute("CREATE TABLE IF NOT EXISTS etlex(id VARCHAR(20), esearch VARCHAR(255), eentry VARCHAR(255), tentry VARCHAR(255), ecat VARCHAR(255), ethai VARCHAR(255), esyn VARCHAR(255), eant VARCHAR(255))")

cursor.execute("INSERT INTO etlex("esearch", "eentry", "tentry", "id", "ecat") VALUES ("a", "a", "หนึ่ง (คํานําหน้าคํานามเพื่อแสดงว่าคํานามนั้นๆ ไม่ชี้เฉพาะ)", "0", "DET");")
cursor.execute("INSERT INTO etlex("esearch", "eentry", "tentry", "id", "ecat") VALUES ("A", "A", "อักษรตัวแรกในภาษาอังกฤษ", "1", "N");")
cursor.execute("INSERT INTO etlex("esearch", "eentry", "tentry", "id", "ecat") VALUES ("a", "a", "อักษรตัวแรกในภาษาอังกฤษ", "2", "N");")
cursor.execute("INSERT INTO etlex("esearch", "eentry", "tentry", "id", "ecat") VALUES ("a posteriori", "a posteriori", "จากผลไปสู่เหตุ", "324", "ADJ");")
cursor.execute("INSERT INTO etlex("esearch", "eentry", "ethai", "tentry", "id", "ecat") VALUES ("A.B.", "A.B.", "ศิลปศาสตรบัณฑิต", "ABBRE", "ABBR");")
cursor.execute("INSERT INTO etlex("esearch", "eentry", "tentry", "id", "ecat") VALUES ("A.D.", "A.D.", "ปีคริสต์ศักราช (คําย่อของ Anno Domini)", "5", "ADJ");")
cursor.execute("INSERT INTO etlex("esearch", "eentry", "ethai", "tentry", "id", "ecat") VALUES ("zymurgy", "zymurgy", "การกลั่นสุรา", "การหมักสุรา", "83231", "N");")
database.commit()
database.close()
```

Fig. 4. Shows example table in SQLite DBMS

Sample source code for Insert Thai English XML Format to SQLite DBMS which is specific like

Thai English Vocaburary for Sqlite DBMS

```python
# -*- encoding: utf-8 -*-
import sqlite3

DB_NAME = '/mnt/sdcard/telex.db'

database = sqlite3.connect(DB_NAME) # Create a database file
cursor = database.cursor() # Create a cursor

cursor.execute("CREATE TABLE IF NOT EXISTS telex(tsearch VARCHAR(255), tentry VARCHAR(255), eentry VARCHAR(255), ethai VARCHAR(255), sant VARCHAR(255))")

cursor.execute("INSERT INTO telex("tsearch", "tentry", "ethai", "tentry", "id") VALUES ("83231", "83231", "ปีคริสต์ศักราช", "ปีคริสต์ศักราช", "83231", "ADJ");")
database.commit()
database.close()
```
V. PROGRAM TO LOOK UP WORDS

The program will show the display and wait for users to put the words to look up. Users can type via keyboard. They can choose between Thai and English or they can use standard keys on android mobile phones. Users press "enter" and the program will look up the words then show the data about the words on the screen of the android mobile phones.

The dictionary program depends on the program to look up words in the database which was converted from XML. The dictionary program depends on the program to look up words in the database which was converted from XML.

The dictionary program depends on the program to look up words in the database which was converted from XML. The dictionary program depends on the program to look up words in the database which was converted from XML. The dictionary program depends on the program to look up words in the database which was converted from XML. The dictionary program depends on the program to look up words in the database which was converted from XML.

VI. TESTING RESULTS

In the test, the word “com” was to be looked up on the
android mobile phone display. After “enter” was pressed, the program would show the list of vocabulary on display and user can click on the checkbox list for the translation on the screen as shown in Figures 6, 7, 8 and 9.

### VII. SUGGESTIONS FOR FURTHER DEVELOPMENT

The words which were not available in the dictionary were normally proper nouns, for example, names of persons or places, new words and words from foreign countries. In the future, the program will be developed to include new words. The translation for sentence can be done if the sentence undergoes Thai word segmentation by using dictionary [15] because this dictionary is to look up only words in the dictionary. In the future, there might be a program which can perform Thai word segmentation by rules [16] to check the rules about Thai language in terms of mixing letters, spacing, and beginning a new paragraph. This research can be further developed by giving sounds for the word which is looked up as well.

### VIII. REFERENCE


Chakkrit Saengkaew is a lecturer at department of Information Science, Faculty of Informatics, Mahasarakham University Thailand. He received the B.Sc. degree in Information Technology from Faculty of Science, King Mongkut's University of Technology Thonburi in 1997. He obtained his M.S. Ind. Ed. degree in Computer and Information Technology from Faculty of Technical Education, King Mongkut's University of Technology Thonburi (KMUTT) Thailand in 2004. During 1997-2004, he works in the field of Computer and Information Technology. He has experience in many positions such as Programmer, System Engineer and Project Manager. He is an author of the book “Python Programming Language by yourself” – best seller computer programming book in Thailand (year 2005).