

Dictionary-based Translation Feature in Open Source LMS ~ A Case Study of Thai LMS : LearnSquare ~

Orrawin Mekpiroon, Pornchai Tammarattananont, Narasak Apitiwongmanit,
Neetiwit Buasroung, Buntita Pravalpruk, Thepchai Supnithi
National Electronics and Computer Technology Center

{*orrawin.mekpiroon, pornchai.tammarattananont, narasak.apitiwongmanit, neetiwit.buasroung, buntita.pravalpruk, thepchai.supnithi*}@nectec.or.th

Abstract

One of the problems on using English in Thai students is vocabulary limitation. This paper proposes a methodology to ease learning curve for English content by adding an English-Thai dictionary based translation features in Thai open source Learning Management System called LearnSquare. Each vocabulary is automatic translated by placing a pointing device over them.

1. Introduction

English is recognized as an important language for communication in cross culture. Based on the ETS TOEFL score statistics [1], the average score of paper-based test for Thai is 500. It is inadequate for studying in English program curriculum, which requires score at least 550.

Thai are proficient in technologies such as electronic applications, biological products. However, English language skills for Thai are below average. This poses difficulty in knowledge sharing of specific resources with other countries [2]. IT valley in Maehongson province is a project under Thai government, which is aimed to develop an IT community in rural area. It is observed that while IT skills can be developed, the same does not apply to English language skills.

Currently, Thailand focuses on developing more local software and Thai language web sites to expose more people to the Internet. E-learning system becomes an important issue to be developed as localized software.

English contents are developed for a long time and contain different styles for supporting wide range of users. Lacking of good Thai contents, English contents have been used intensively in some respectable high schools and colleges. In this internet era, knowledge is shared abundantly. Therefore, e-learning system is required to add an English-Thai dictionary based translation features to help learning English content. A dictionary based translation will help Thais to reduce the task on looking up dictionary and understanding vocabulary concept.

2. Background Component

LearnSquare is a Thai open source LMS which is supported by National Electronics and Computer Technology Center (NECTEC), Thailand [3]. LearnSquare is a module based system which supports platform independence feature. Content is represented in IFRAME tag, which enables users to design and implement easily. Currently LearnSquare is widely used in a lot of learning institutes, organizations and private company in Thailand. Moreover, LearnSquare becomes a fundamental system involved in a lot of projects related to schools in rural area that face the poor quality teaching problems.

LEXiTRON is an online dictionary developed by NECTEC since 2003 [4]. The dictionary was originally constructed from a corpus which consists of frequently-used vocabularies in many topics from reliable publications. Currently, the database has more than 53,000 of English and more than 35,000 of Thai [5]. We apply data in LEXiTRON which is open freely to public users in our work.

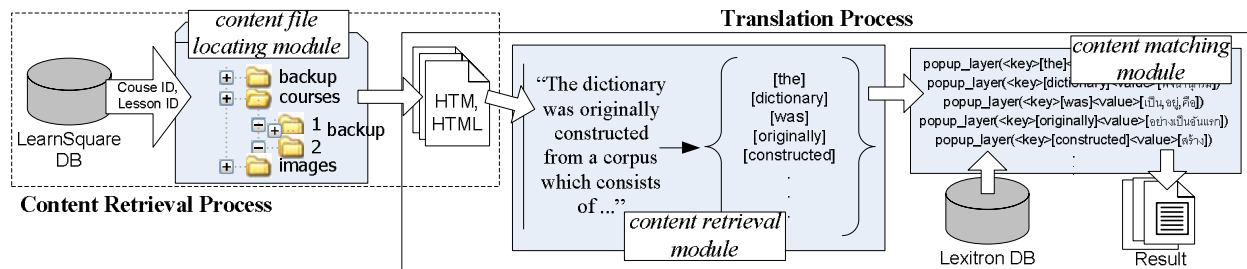


Figure 1 LearnSquare System Architecture with Dictionary Based Feature.



Figure 2 Output screenshot of vocabulary translation.

3. System Architecture

Figure 1 shows LearnSquare System Architecture with addition of an English-Thai dictionary based translation feature. The system consists of two databases; LearnSquare database which stores contents and general information for users and LEXiTRON database which stores vocabularies and their meanings.

A dictionary based translation feature composes of three modules; *content file locating module* which is a module for detecting correct file location, *content retrieval module* which is a module for separating content and tag form, *content matching module* which is a module for matching content and LEXiTRON database and visualize results in popup layer. All data represented in each module uses HTML style.

4. System Process

There are two main processes in this system as shown in Figure 1. *Content retrieval Process* detects the content file location. Content is developed and stored in LearnSquare at “course” folder which can be detected by getting course ID and lesson ID from database. After file location are detected, it will be sent to *Translation Process*. Since there are varieties types of data representation, we focuses only HTML format, which is mostly used in LearnSquare.

Translation Process starts with separating content and HTML format. Content is segmented into words by using space. Each word is matched with

vocabularies in LEXiTRON database. Translation meanings are retrieved to display in popup layer.

The output screenshot is shown in Figure 2. When a student clicks to choose a lesson, it will be delivered to him/her with translation meaning. Student can translate vocabularies by placing the mouse over the desired words.

5. Conclusion and Future Work

In this paper, we developed a dictionary based English-Thai translation feature and integrated into LearnSquare. Users tentatively understand English lessons in less time. Vocabularies can be easily translated by navigating the pointing device over the desired words.

However, this function is limited to web based content fundamental form of HTML. We plan to extend features to cope with other file type. Moreover, we concern with using Machine Translation (MT) to integrate with dictionary for a better result.

6. References

- [1] EST, *TOEFL Test and Score Data Summary for TOEFL Computer-Based and Paper-Based Tests: 2005-2006 Test Year Data*, 2007.
- [2] Wiriyachitra, A., “English Language Teaching and Learning in Thailand in this Decade”, *ThaiTESOL Focus*, Thailand, 2002.
- [3] LearnSquare. <http://www.learnsquare.com>.
- [4] LEXiTRON. <http://lexitron.nectec.or.th>.
- [5] Kanokorn Trakultaweekoon, Peerachet Porkaew and Thepchai Supnithi, “LEXiTRON Vocabulary Suggestion System with Recommendation and Vote Mechanism”, *SNLP2007 The Seventh International Symposium on Natural Language Processing*, Thailand, Dec 2007, pp. 43-48.