A Case Study of the Localization of an Intelligent Tutoring System

Phaedra Mohammed and Permanand Mohan

Department of Computing and Information Technology, The University of the West Indies, St. Augustine, Trinidad and Tobago
{phaedra.mohammed@gmail.com, pmohan@tstt.net.tt}

Abstract. The matter of a one-size-fits-all approach towards the development of culturally-relevant educational software is debated with one side arguing for internationalization and the other side arguing for localization. This paper takes a pragmatic look at the issues involved in localization and aims to shed light on the strengths and limitations of undertaking culture as a design feature. With an emphasis on the application layer, the paper investigates the requirements and steps that need to be taken when using cultural contexts in educational software. It describes the design of a localized intelligent tutoring system developed for the context of Trinidad and Tobago and discusses how the prototype was evaluated in two separate studies which looked at learning gains, students' opinions, attitudes, and preferences for localization.

Keywords: Localization, cultural translation, intelligent tutoring systems

1. Introduction

This paper is set in the context of Trinidad and Tobago. With a GDP of $20,400US [2] Trinidad and Tobago is one of the wealthiest nations in the Caribbean. Liquefied natural gas, petroleum and its byproducts make up the bulk of the country's exports and account for approximately 40% of the country's GDP [7]. The average population size is 1.2 million, life expectancy is estimated at around 72 years, and literacy rates are over 98% for ages 15 and older [2]. The country is becoming more modernized as evidenced by the increasing number of Internet users (growth from 8% in 2000 to 48% of the population in 2012) and large number of cell phone users (over 1.8 million) [2]. Although access to personal computers is not as widespread with roughly less than 20% of the population having access, the government of the country provided free laptops to entry-level students in secondary schools in 2010. The challenge that now arises for the country's education sector is whether the software on these machines can support learning in the context of Trinidad and Tobago.

Accommodating for learner diversity based on cultural backgrounds is becoming a major personalisation focus with the increasing drive towards globalization as evidenced in Trinidad with the distribution of laptops. Despite this drive, the knowledge and processes for incorporating culture have not been clearly defined with automation

in mind. The matter of a one-size-fits-all approach towards the development of educational software is debated with one side arguing for internationalization and the other side arguing for localization. Proponents of localized approaches argue that culture increases the credibility, realism, familiarity and acceptance of educational systems with the end result being a higher quality learning experience [4, 5]. On the other hand, techniques have been described for creating internationalized designs that do not target a particular culture and avoid cultural specificity but still cater for the needs of a learner. This viewpoint is based on arguments that cultural designs tend to be cosmetic and stereotypical, suffer from designer bias [1], and are overall difficult to automate [8].

The purpose of this paper is not to take a particular side or argue for or against a particular approach. Rather, this paper takes a pragmatic look at the issues involved, and aims to shed light on the strengths and limitations of undertaking culture as a design feature in educational systems. It focuses on the requirements and steps involved in carrying out localization at the application layer. The paper then investigates the requirements and steps that need to be taken in order to use the cultural context of student in educational software and then describes the design of a localized intelligent tutoring system (ITS) developed for the context of Trinidad and Tobago with these requirements in mind. Next, the paper discusses how the system was evaluated in two separate studies which looked at learning gains, students' opinions and attitudes towards the system, and student preferences for localization. The paper concludes with an outline of the lessons learned from these studies and potential research directions for localized systems.

2. The Strengths and Challenges of Localization

Educational frameworks and systems from different sources are typically repurposed in order to reduce costs associated with content development, to replicate proven results with learning gains, and to set standards in educational curricula. Repurposing of such environments entails some form of localization since the design of user interfaces, the selection of teaching strategies, the format and content of the educational material all vary depending on the contextual background of the developers. In order to relate to students and in some cases not offend others, localization of educational environments has been recommended and there have been benefits cited in the literature for students such as increased motivation levels [3, 6].

There are several challenges associated with localization such that actual systems are limited in practice. Many developers have shied away due to the complexity in reliably representing aspects of a particular culture and because of the ill-defined nature of culture [1, 8]. The costs can be higher in the long run because of the amount of effort required in cataloguing cultural knowledge. In addition, localization requires many pieces of metadata for adaptation to go beyond keyword insertion and colour changes. Learning gains seen in one country are not guaranteed in another since many variables exist across countries such as the system may not be used in the same way as in the original country, there might be limited internet connectivity, different hard-

ware, untrained teachers, or even a different curriculum. Lastly, a localized solution runs the same risk of being as offensive or irrelevant as the original version since a student does not belong to a discrete cultural group, and more than just cultural knowledge is required in order to relate to students; a sound understanding of environmental context is critical.

3. Requirements for Effective Localization at the Application Layer

There are three entry points for localization in educational software systems: the presentation layer, the application layer, and the data layer. The application layer selects, modifies, aggregates, and formats raw content based on the user's input, history of events and intended instructional goals. Firstly, data is needed about the student's country of residence (target country) because this data defines the environmental context that the student is familiar with. Such data is typically available from the country's national statistical office or department and can be used to set the scope of the localization. For example, statistics on population density, economic activity, and religious group distribution can be used. Secondly, demographic data from the student is required in order to model his/her contextual background and to know what type of localization to carry out and to what extent. This data comes from users directly entering information about themselves in a form for instance and can include a user's age, gender, native language and residence location for instance. Thirdly, data is required on the contextual groups in the target country specifically what kinds of contextual elements are familiar, appropriate, and relevant to a specific group. This is needed in order to control the direction of the localization process. Next, language rules and localized natural language terms are needed in order to translate textual content and adapt the cultural meaning of the content to suit the student. Semantic markup on these cultural terms is required in order to know how to use the terms in sentences correctly and also to be able to interpret image context in a culturally appropriate manner. Lastly, real-time localization is required in order to adapt content to suit different student preferences.

4. Localized System Design

This section describes the design of a contextually-relevant Intelligent Tutoring System (ITS). A modular design was chosen because of the complexity involved in delivering culturally-relevant instructional experiences. Flexible alteration and improvement of the component features are easily accommodated as a result. Many of the components featured in the design of the localized web-based ITS are based on traditional ITS components but have been modified to include localized data and functionality. A cultural student model, a pedagogical module, cultural heuristics and, a content repository make up the major architectural units of the design.

The cultural student model records the pedagogical events, performance-related student data, suggested hints and instructional guidance given to the student. Cultural

background data is assimilated into this model and is made up of demographic data and cultural influences. The pedagogical module consists of instructional rules that constantly access and update the student model in response to input data from the student and events captured from the screen. They control how and when instructional feedback is given, and they manage the selection and transition of learning activities.

A culturally-relevant instructional approach requires that cultural references made by software systems should be applicable to the learning content, familiar to the students, authentically rendered, and integrated into the context of the instructional material [4, 5]. Cultural rules modify the textual portions of instructional content such as question descriptions, scenarios, hints, and instructional feedback produced by systems. The localization of visual portions of these systems, namely the multimedia related to the learning activities, is also handled by these rules. Textual modifications include customizing the language of the instructional feedback, whereas the multimedia localization involves swapping in cultural assets for generalized assets such as images. The scope of the languages used for localization in this paper is restricted to English-based Creole languages. These languages are useful for educational environments because they foster comfortable learning experiences especially in domains that are seen by students as problematic or difficult to cope with [6]. Textual outputs of the localization process are therefore sentences expressed in a local dialect specifically mesolect forms[1], and feature equivalent cultural lexical terms. In order for these rules to properly localize the learning activities and determine appropriate feedback for the student, the learning materials need to be accompanied with metadata descriptions that identify content that can be localized such as hints, question descriptions, and images.

The content repository handles the organisation and distribution of all instructional and cultural assets to a content aggregator. Localized ITSs rely on reusable content more than non-cultural ITSs because of the additional dimension of cultural personalisation; this was the basis for having a separate asset repository - reusability. The content repository primarily hosts all of the educational and interface-related material used by the system and the student model. For example, multimedia files related to the interface's look and feel, such as icons, logos, and those related to the learning exercises (scenario pictures, feedback pictures) are stored here together with educational material such as question descriptions, solutions, feedback files, and topic hierarchies. Each of these assets is described by their asset metadata descriptions which define the context of use and the nature of the assets. These descriptions are indispensable in the design because they facilitate reuse and exchange of compatible assets. Both the pedagogical module and cultural heuristic component use these descriptions when making instructional and localization-related decisions.

5. System Implementation

Two web-based ITSs were implemented based on the software architecture described in the previous section. One system was localized for Trinidad and Tobago's

[1] A variety of language in a Creole continuum that is intermediate between the standard form (acrolect) and forms that diverge greatly from the standard form (basilect).

context (Culturally Relevant Instructional Programming System – CRIPSY) while the other remained generic (Instructional Non-CulturAl Programming System – INCAPS).

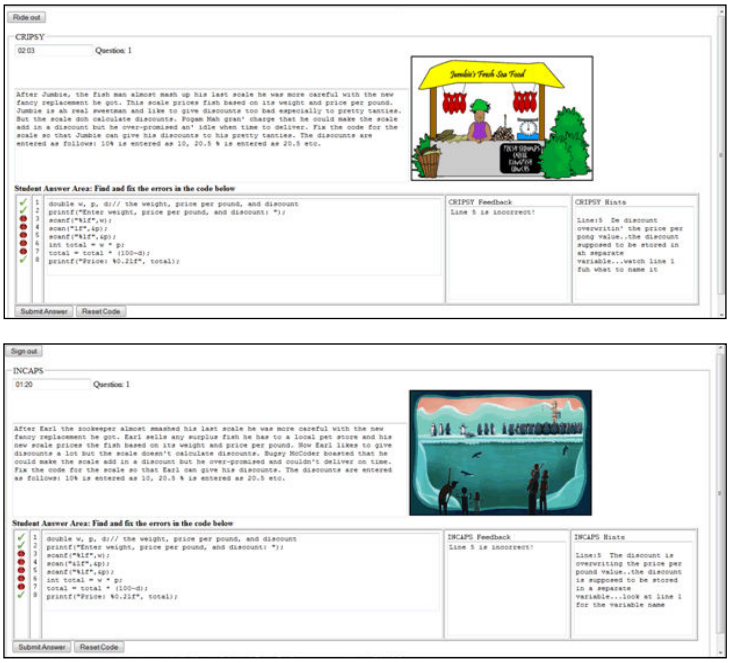

Fig.1. Screenshots of CRIPSY (top) and INCAPS (bottom) featuring localized and non-cultural versions of the same programming exercise.

Both systems were built for the same educational domain (Computer Science programming) were identical from a functional standpoint, and were implemented using Java-based tools and technology which facilitated seamless integration of the various components into complete systems. The pedagogical module, student model, and cultural heuristic component were implemented using JESS (Java Expert System Shell) rule engines and rules. At the data level, simple file formats were used to manage the content repository since rule engines handle data manipulation primarily using facts. Constructivism and situated cognition were selected as the major instructional strategies since analytical programming skills were being targeted and also because of the good fit between situated cognition and culturally-aware instruction. The development of localized assets for CRIPSY was done semi-automatically and manually. The programming exercise descriptions, parts of the exercise code and instructional hints were localized using subtle, careful use of cultural semiotics, specifically cultural names of objects and foods. As shown in the screenshots in Figure 1 above, both systems used the same instructional content but differed in the expression of the content, that is, cultural and non-cultural. A minimalist interface design was used and instructional feedback consisted of identification of correct/incorrect lines of code, hints for the incorrect lines, and general informative guidance.

6. Prototype Evaluation and Results

A previous study was done using both prototypes and participant details are reported in [6]. That study revealed that the use of both systems resulted in significant increases in student test scores and one design issue affected students in particular: the language used in the localized system. The density of the localisation distorted the system's content into basilect[2] Creole which for some students made the sentences difficult to read quickly and therefore difficult to understand quickly. This observation stimulated the consideration of incorporating a customizable language density scale in the localized system so that students may adjust the language to suit their own preferences. The desirability of such a feature was assessed in the second study.

The second study aimed to find out which system appealed more to the students, what kinds of conditions would impact if at all upon their preferences, and whether a language localization slider would be a desirable feature and why. Fifty-eight (58) students from the previous study participated in this study. The control and test groups were switched so that students would have used both systems at the end of this study. A similar procedure was followed where each student's username activated their newly assigned system and the systems timed out after 30 minutes. A short questionnaire was administered and the results are shown in Table 1 below. After using both systems, more students (68.9%) preferred the localized system (CRIPSY) over the non-localized one (INCAPS). The majority of students reported that they wanted the localization slider (79.3%) while the minority did not see the slider as necessary. Increasing question difficulty did not seem to influence student preference for either system. Lastly, more students (13.8%) changed their preferences from the localized system to the non-localized system when server glitches or software problems occurred compared to those whose preferences were reversed in favor of the localized system (3.4%).

Feedback Topic	**Student Preferences**	**Percentage of Students**
Localization Slider	Wanted localization slider	79.3
	Did not want localization slider	20.7
System of Choice	Preferred localized system in general	68.9
	Preferred non-localized system in general	31.1
Preference in Relation to Question Difficulty	Preferred localized system	48.3
	Preferred non-localized system	48.3
	No preference/no response	3.4
Preference change if glitches occurred in System of Choice	Changed to non-localized from localized	13.8
	Changed to localized from non-localized	3.4
	No change in preference	3.4
	No response	79.4

Table 1. Percentage of student preferences categorized by feedback topics

[2] The basilect variety of English-based Creoles is the most distorted from Standard English.

7. Analysis of Results

The results reported in the previous study using the localized ITS, CRIPSY, confirmed the assumption that cultural interventions do indeed have positive effects on learners and provide empirical evidence in support of localized learning systems. Overall, the students liked the localized system primarily because of the reasons outlined in the earlier studies in [6] namely enriching learning experiences and humour. The use of culture created a familiar setting and it was done in a way that was interesting to the students. A larger percentage of students rated the programming exercises as easier and the instructional feedback/hints as more helpful for the localized system although similar guidance was given in the control system.

In the second study, the majority of students (79.3%) wanted the localization slider. The most common reasons given by students for wanting a slider included: wanting control over the timing and the degree of localization, wanting to change the localization to suit their moods, and wanting to be able to explore the different degrees of localization. An interesting trend in the responses of the students who did not want the slider was their dislike of the localization. Many of them stated that the localization resulted in descriptions that were longer to read and had too much localized language which was confusing. Longer descriptions were indeed the case since the highest density of localization was used in generating content for the study, and therefore the maximum number of lexical insertions and replacements possible for the content was made. This indicates that there is a strong need for the slider since the students essentially wanted to choose their own levels of localisation. Another interesting result is the low student tolerance for software faults in the localized system evidenced by the larger numbers of students who changed their preference to the non-localized system when faults occurred in the localized one. A possible cause could be that students perceived the localized system as being inferior because of the cultural behaviour of the system in using language levels (basilect) typically associated with the less educated in Trinidad and Tobago. This implies that being able to dynamically adjust localization is crucial for system acceptance by students.

8. Conclusion and Future Research

Culture is rapidly becoming an important consideration in the design of educational software firstly because of the increase in the number of users accessing software over the Internet, and secondly because of the sheer diversity in the cultural backgrounds of these users. Conventional learning has often taken place in a localized setting with a teacher guiding one or more students in their search for knowledge. With the advent of the Internet, this traditional setting has changed drastically since students now have access to teachers and educational material from over wide distances. Consequently, these students are exposed to a variety of educational tools, teaching strategies and learning materials which were not developed with their own personal needs in mind. This has dramatic usability implications especially when the mainstream culture for which e-Learning materials are designed clashes with that of the users.

Based on the encouraging evidence established by these studies, the research discussed in this paper demonstrates a practical approach towards developing a localized web-based learning environment. By leveraging research from various fields such as Intelligent Tutoring Systems and culturally-aware instruction, this research shows how some of the complexity of localization can be managed and how aspects of intelligent tutoring can be localized. Empirical evidence indicates that localized systems perform as well as traditional tutoring systems and are potentially superior at creating relaxed, engaging learning atmospheres for the Computer Science programming domain. However care must be taken to ensure that the cultural enhancements match the tolerance level of the student users. Further refinement and improvements are planned for the systems described. A limited amount of cultural automation was undertaken, so expansion of the cultural coverage is necessary. Additional features such as deeper cultural student profiling, adjustable language density and greater tutoring flexibility are also part of the plans intended for this research.

References

1. Blanchard, E.G., Mizoguchi, R., Lajoie, S. P.: Structuring the cultural domain with an upper ontology of culture. In Blanchard, E., Allard, D. (Eds.): The Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models. 179-212. Hershey, PA: IGI Global. (2011)
2. CIA The World Factbook. Available online: [https://www.cia.gov/library/publications/the-world-factbook/geos/td.html] (2012)
3. Gilbert, J.E., Arbuthnot, K., Hood, S., Grant, M. M., West, M.L., McMilllan, Y., Cross, E.V., Williams, P., Eugene, W.: Teaching Algebra Using Culturally Relevant Virtual Instructors. The International Journal of Virtual Reality 7, 1, 21-30 (2008)
4. Henderson, L. Theorizing a multiple cultures instructional design model for e-Learning and teaching. In Edmundson A. (Ed.): Globalized E-Learning Cultural Challenges. 130- 154. London, UK: Information Science Publishing. (2007)
5. McLoughlin, C.: Adapting e-Learning across cultural boundaries: A framework for quality learning, pedagogy, and interaction. In Edmundson A. (Ed.): Globalized E-Learning Cultural Challenges. 223- 238. London, UK: Information Science Publishing. (2007)
6. Mohammed, P., Mohan, P.: The design and implementation of an enculturated web-based intelligent tutoring system for Computer Science education. In I. Aedo, N-S. Chen, D.G. Sampson, J.M. Spector, Kinshuk (Eds.): 11th IEEE International Conference on Advanced Learning Technologies (ICALT). 501-505. Washington, USA: IEEE Computer Society. (2011)
7. Trinidad and Tobago. Available online: [http://en.wikipedia.org/wiki/Trinidad_and_Tobago] (2013)
8. Rehm, M.: Developing enculturated agents: Pitfalls and strategies. In Blanchard, E., Allard, D. (Eds.): The Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models. 362-386. Hershey, PA: IGI Global. (2011)