Diagnosing Language Transfer in a Web-based ICALL that Self-Improves its Student Modeler

Victoria Tsiriga  
Department of Informatics,  
University of Piraeus,  
80 Karaoli & Dimitriou St.  
Piraeus 18534, Greece  
vtsir@unipi.gr

Maria Virvou  
Department of Informatics,  
University of Piraeus,  
80 Karaoli & Dimitriou St.  
Piraeus 18534, Greece  
mvirvou@unipi.gr

Abstract
This paper describes a Web-based Intelligent Computer Assisted Language Learning (ICALL) system that may self-improve its student modeler while it acquires knowledge from its users. The ICALL system is called Web-PVT and is a Web-based language tutor for the domain of the passive voice of the English language. The student modeling approach of Web-PVT considers the issue of language transfer of primary importance and exploits the fact that Web-based systems have a large and heterogeneous number of users. In particular, a machine learning reasoning mechanism is used, which is based on recognized similarities between users in order to predict the proneness of a new student to make mistakes of a particular type. The similarity between two students is calculated taking into account their mother tongue as well as other languages that they may already know. Thus, error proneness due to language transfer may be modeled.

1. Introduction
Web-based educational systems aim at reaching a heterogeneous group of learners in settings where no teacher may be available to help users during the learning process. Therefore, such systems have to be highly adaptive to serve the needs of students of such diverse backgrounds. To be adaptive, a Web-based educational system should be able to draw inferences about individual students. Such inferences may concern the students’ knowledge level, misconceptions, and certain other students’ attributes that may play a significant role in the students’ performance to exercises and/or way of learning. Especially for Web-based Intelligent Computer Assisted Language Learning (ICALL) systems, a characteristic that would be of great importance would be the student’s prior knowledge of other languages (such as his/her mother tongue). This stems from the fact that students’ performance in language learning is greatly influenced by the issue of language transfer. Language transfer is the interference resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired [1]. Indeed, a major issue in the error analysis of ICALL systems is the native language of the student. In particular prior to the development of such systems, empirical studies are conducted and a list of the most likely errors is identified, based on the mother tongue of the target student population (e.g. [2]). However, a Web-based system must adopt a more general scheme in order to accommodate the international nature of the Internet and cases where the native language of a user might not be known [3]. We argue that using a machine learning reasoning mechanism would allow an ICALL system to learn how a specific language may interfere in the process of the student’s learning the target language. In particular, the system could use information from existing groups of students who share the same mother tongue and perhaps other languages they know, in order to make inferences about a new student who belongs to one of these groups.

In this paper we describe an approach to model the error proneness of students, with respect to the issue of language transfer, in Web-based ICALL systems. We have incorporated our approach in an adaptive and intelligent language tutor for the domain of the passive voice of the English language. The system is called Web Passive Voice Tutor (Web-PVT) and is delivered via the
WWW. Previous versions of Web-PVT [4] did not have the ability to self-improve the student modeler of the system based on expertise acquired through the use of the system.

In particular, the current version of Web-PVT models the language transfer of a new student based on information derived from the interaction performance of other students. These students are similar with the new student with respect to their mother tongue as well as other languages that they may already know. However, they have already used Web-PVT sufficiently for the system to have been able to construct their individual student models inferred from extensive observation of their own behavior.

2. Overview of Web-PVT

Web-PVT incorporates techniques from Intelligent Tutoring Systems (ITS) and Adaptive Hypermedia (AH) to tailor instruction and feedback to each individual student. In particular, Web-PVT has a student modeling component that models the students’ level of knowledge and possible misconceptions. The students’ level of knowledge and error proneness in certain concepts are used by the system to support the students while they navigate through the course material. Web-PVT uses a combination of two link adaptation techniques to help the student while navigating through the structured theory hyperdocument; namely adaptive link annotation and direct guidance.

The student model is also utilized in order to perform error diagnosis and advice generation in cases of error when solving exercises. When the system performs error diagnosis, it tries to attribute a student’s mistake to a certain category of error and find the cause if this error (e.g. carelessness, language transfer, etc.). In some cases a mistake of a student may be attributed to more than one categories of error. In cases like this the system takes into account the individual characteristics of the student, that have been recorded in previous interactions, in order to resolve the ambiguity and formulate an appropriate piece of advice.

However, when a student interacts with Web-PVT for the first time, individual characteristics concerning the error proneness of the student are not available. In cases like this, Web-PVT tries to predict the error proneness of students concerning a number of predefined categories of error (described in the next section). The kinds of error a student makes is greatly influenced by the mother tongue of the student and/or foreign languages s/he may be learning. Furthermore, the proficiency level of the student in the domain being taught also plays a significant role in the student’s proneness to make mistakes of a particular type, irrespective of the native language of the student and/or the foreign languages s/he may be learning.

Indeed, intermediate students who have Greek as their mother tongue may type “the police has arrested him” instead of “the police have arrested him” due to language transfer. However, in case of an advanced student, this error is not expected to be made with the same frequency.

Therefore, being able to predict the error proneness of a student due to language transfer, requires the system to combine information that concerns the knowledge level of the student, her/his mother tongue as well as other languages s/he may already know. Such predictions are usually made using predefined assumptions that result from extensive empirical studies. However, conducting empirical studies about students with diverse native languages would be extremely time-consuming. Furthermore, the predictions made using such an approach are static and cannot be updated unless a human does so explicitly. An alternative approach that aims at overcoming these drawbacks may incorporate a machine learning approach to allow the system self-improve its predictions.

In Web-PVT, we use the distance weighted k-nearest neighbor algorithm to make predictions about the student’s proneness to make mistakes of a particular category. In general, nearest neighbor learning algorithms typically store all of the available training examples during learning. These algorithms use a distance function to determine how close a new query instance is to each stored instance, and use the nearest instance or instances to classify the query instance [5]. The basic idea of the distance weighted k-nearest neighbor algorithm [6, 7, 8] is to weigh the contribution of each of the k neighbors according to their distance to the query point, giving greater weight to closer neighbors [9].

3. Error Categories and Language Transfer

Prior to the development of the system we conducted an empirical study that involved human teachers and students. Based on tests set to students, common mistakes that students frequently made while learning the passive voice of the English language were identified. In particular, human teachers classified common mistakes into nine broad categories of error that may be recognized by the system. These errors are associated with the conversion of passive into active voice and vice versa and also with the prerequisite grammatical concepts, such as irregular verbs. The categories of error are presented in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Error Categories</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Language Transfer</td>
<td></td>
</tr>
<tr>
<td>2. Carelessness</td>
<td></td>
</tr>
<tr>
<td>3. Context transfer</td>
<td></td>
</tr>
<tr>
<td>4. Word choice</td>
<td></td>
</tr>
<tr>
<td>5. Grammar</td>
<td></td>
</tr>
<tr>
<td>6. Syntax</td>
<td></td>
</tr>
<tr>
<td>7. Spelling</td>
<td></td>
</tr>
<tr>
<td>8. Punctuation</td>
<td></td>
</tr>
<tr>
<td>9. Other</td>
<td></td>
</tr>
</tbody>
</table>

Each of the categories of error may be associated with an explanation about the possible underlying cause of the mistake. In Web-PVT explanations have been based on identified strategies that learners may use in order to simplify the task of learning a second language [10]:

1. Language Transfer.
2. Overgeneralization of the target language rules.
3. Ignorance of rule restrictions.
4. Incomplete application of rules.
5. False concepts hypothesized.
6. Carelessness.

Table 1. Categories of the most common students’ mistakes.

<table>
<thead>
<tr>
<th>Category of Error</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accidental slips</td>
<td>The student may have deleted some words and/or may have forgotten to complete the sentence.</td>
</tr>
<tr>
<td>Spelling mistakes</td>
<td>The student may have typed “mather” instead of “mother”.</td>
</tr>
<tr>
<td>Article mistakes</td>
<td>The student may have typed “an book” instead of “a book”.</td>
</tr>
<tr>
<td>Irregular verb mistakes</td>
<td>The student may have used the usual ending “ed” to create the past participle of an irregular verb.</td>
</tr>
<tr>
<td>Absence or redundant presence of the agent</td>
<td>The student may have provided the answer “he is said by people to be wealthy” instead of “he is said to be wealthy”.</td>
</tr>
<tr>
<td>Mistakes in the word that connects with the agent</td>
<td>The student may have written “she was killed by a knife” instead of “she was killed with a knife”.</td>
</tr>
<tr>
<td>Verb tense conversion mistakes</td>
<td>The student may have typed “is cooking” instead of “is being cooked”.</td>
</tr>
<tr>
<td>Singular/Plural mistakes</td>
<td>The student may have entered “English are spoken all over the world” instead of “English is spoken all over the world”.</td>
</tr>
<tr>
<td>Errors in the special cases of the passive voice</td>
<td>The student may have provided the answer “I was let to go out” instead of “I was allowed to go out”.</td>
</tr>
</tbody>
</table>

Language transfer may certainly be a cause of many mistakes. For example, if a student types: “English are spoken all over the world” instead of “English is spoken all over the world”, then s/he has made a singular/plural mistake. This particular mistake may be attributed to language transfer for Greek students because “English” in Greek is always in plural. Another example could be the case where the student has made a spelling mistake, such as “informations” instead of “information”. This error may be attributed either to ignorance of rule restrictions or language transfer. If the student who has made this particular mistake is considered advanced and has French as her/his mother tongue, the most plausible cause of the error would be language transfer, since the noun “information” has a plural form in French.

However, knowing exactly which categories of error may be attributed to language transfer would require eliciting the expertise of many human experts. Each of these experts would ideally specialize on teaching English to students of a particular background (e.g. students having a particular mother tongue, students having a particular mother tongue and learning the same foreign languages, etc.). Therefore, in Web-PVT, the association of the above mentioned categories of error with the underlying cause of language transfer is performed dynamically based on the expertise acquired by the system.

4. Using the Distance Weighted k-Nearest Neighbor to Predict Error Proneness of Students

In order to be able to draw inferences about a student, the system should acquire initial information about this student when s/he interacts with the Web-based ICALL for the first time. For some data concerning the student, direct provision by the student is the only possible source of information. Such information is acquired by Web-PVT using a set of questions posed to the student when s/he uses the system for the first time. In particular, the student is asked to provide information concerning her/his name, her/his mother tongue and other languages s/he already knows. Furthermore, the student is asked to self-categorize herself/himself concerning how careful s/he is when solving exercises. There are three distinct categories concerning the degree of carefulness of a student, namely careless, averagely careful and very careful.

Directly asking students to provide information about themselves is one of the most obvious methods of acquiring information. However, self-assessment is error-prone, since users are often not correctly aware of their own capabilities [11, 12]. Therefore, similarly with ELM-ART II [13], Web-PVT also uses a preliminary test in order to assess the knowledge level of the student in the domain of the passive voice of the English language. In particular, based on the student’s performance on the preliminary test, Web-PVT assigns the student to a category concerning her/his knowledge level. There are four distinct categories concerning the knowledge level of the student, namely novice, beginner, intermediate and advanced.
When the necessary information about the student has been acquired, the system represents this information in a feature vector. The characteristics contained in this vector include the name of the student, the category that s/he belongs concerning her/his knowledge level, the estimation of how careful the student is while solving exercises, her/his mother tongue, as well as other languages that the student already knows. Therefore, this vector is defined as follows:

\(<\text{Student Code}, \text{Name}, \text{Knowledge Level}, \text{Carefulness}, \text{Mother Tongue}, \text{Language1}, \text{Language2}, \ldots>\> \\

In Web-PVT, the distance weighted k-nearest neighbor algorithm is used to make estimations concerning the student’s proneness to make each category of the most common students’ errors (described in the previous section). This is done using information about other students that belong to the same knowledge level category and speak and/or learn the same languages as the new student. However, these students have already used Web-PVT sufficiently for the system to have been able to construct their individual student models inferred from extensive observations of their behavior. The contribution of each of the neighbor students is weighted based on her/his distance from the new student. The distance between students is calculated taking into account the students’ mother tongue as well as their prior knowledge of other languages.

In particular, the distance function used by Web-PVT to calculate the distance between two values \(x\) and \(y\) of a given attribute \(a\) is defined in (1).

\[
d_a(x, y) = \begin{cases} 
1, & \text{if } x \text{ or } y \text{ is unknown, } \\
\text{overlap}(x, y), & \text{else}
\end{cases}
\] (1)

Furthermore, the \(\text{overlap}\) function is defined as shown in (2).

\[
\text{overlap}(x, y) = \begin{cases} 
0, & \text{if } x = y \\
1, & \text{otherwise}
\end{cases}
\] (2)

The overall difference measure of two students \(s_x\) and \(s_y\) is then calculated as defined in (3), where \(n\) is the number of student characteristics that are used to measure the distance between two students. In the case of Web-PVT these characteristics are the estimation of how careful a student is while solving exercises, her/his mother tongue, as well as the foreign languages the student already knows.

\[
\Delta(s_x, s_y) = \sum_{a=1}^{n} d_a(x_a, y_a)
\] (3)

Another important decision that should be made when applying the distance weighted k-nearest neighbor algorithm is the definition of the number of neighbors that would participate in the classification task. In our approach, we have set the number of \(k\) to be the number of students that belong to the same category, concerning their knowledge level, with the student in question. This decision has been based on the fact that students that belong to different knowledge level categories make mistakes of each category in a different percentage. Indeed student mistakes may be attributed to language transfer, but they may also be misconceptions that stem from lack of knowledge of the student. For example, intermediate and advanced students may both not make mistakes in verb tense transformation, since this concept is simple. However, when solving exercises that evaluate special cases of the passive voice, an advanced student makes less mistakes when compared to an intermediate student.

Finally, a classification function has to be chosen that uses as input the instance that should be classified (new student) and the k-nearest neighbors (students that belong to the same category concerning their knowledge level). In the case of Web-PVT, the system makes estimations about the student’s proneness to make each category of the most common students’ mistakes, based on the known values of these attributes that are acquired by students that have already been registered in Web-PVT. In order to predict the new student’s error proneness \((s_q)\), we use a distance weighted mean value of the error proneness of the \(k\) students that belong to the same knowledge level category with the new student \((s_1, s_2, \ldots, s_k)\). The vote of each of the neighboring students is weighted according to the inverse square of its distance from \(s_q\). Therefore, for each category of the most common students’ mistakes \((\text{Category}_x)\), the function that predicts the proneness of the new student to make mistakes of this category is presented in (4).

\[
\text{Error}(	ext{Category}_x, s_q) \leftarrow \frac{\sum_{i=1}^{k} w_i \text{ Error}(	ext{Category}_x, s_i)}{\sum_{i=1}^{k} w_i}
\] (4)

where \(w_i\) is the weight of the contribution of each student and is calculated using (5).
\[ w_i = \frac{1}{\Delta(s_q, s_i)^2} \] (5)

To accommodate the case where the query student \( s_q \) matches exactly one of the students \( s_i \) that are used as training examples and the denominator \( \Delta(s_q, s_i) \) is therefore zero, we assign \( w_i \) to be equal to 1 (maximum weight) in this case.

5. Conclusions

In this paper we have described the approach taken by Web-PVT to model the error proneness of a new student taking into account the important issue of language transfer. The importance of this issue stems from the fact that students often use already acquired knowledge while learning a new subject. Our approach to student modeling is based on recognized similarities of the new student with other students that have already interacted with the system. In Web-PVT the similarity is estimated taking into account the students’ mother tongue, as well as their prior knowledge of other languages.

6. References