International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 01 | Jan-2018 www.irjet.net p-ISSN: 2395-0072

Language Translation for E-learning Systems

Ruchika Sinhal[1],Kapil Gupta[1] & Anuj Jain[2]

[1]Dept of CSE, DMIETR, [2]SE, Infosys, Pune

Abstract - *ICT as a medium for teaching is becoming more and more acknowledged. In this article we wish to share some views for translation of e-learning material for stimulating growth of both students and the teacher. The targeted approach of E- learning is to give methodical, systematic and efficient outcomes in learning. This can be accomplished by enhanced communicational arrangement between students and teachers. It may result in reduced learning time and cost of teachers' training programs. Moreover, E-learning provides for efficient use of resources compared to traditional methods. On the other hand, diversified and poorly organized abundant information in different languages is a drawback of E-learning. To overcome this stumbling block, machine translation techniques are recommended to find useful solution in generating study- materials in native languages understandable by the students. Machine Translation and learning is a highly useful, efficient and beneficial tool for providing specific ways to make productive and constructive decisions in teaching methods and conveyance models. Our focus will be on what are different machine translation approaches and types for supporting student-centered learning, increasing student motivation, individualization and cooperation in translating e-learning material, at the same time developing a feeling of "us" and of belonging together in diverse environment.*

Keywords: Benefits of E-learning, ICT, Language learning, Machine Translation types, Machine Translation approaches

1. Introduction

E-learning tends to create dissenting opinions. Some educationalists appreciate its values; others tend to be rather reserved to the option of having the electronic environment "overtake the classroom". Over the years, Information and Communications Technology (ICT) developed rapidly to give rise to new services out of which online learning evolved. We are dependent on the World Wide Web (WWW) for delivering information. The Web 2.0 standard made content delivery systems presentable in a way that enabled learners to better perceive concepts explained by Clark, R. C. & Mayer, R. E. in [1]. After the development of Web 2.0, participative and interactive learning could become feasible explained by Greenhow *et.al* in [2]. Today several E-leaning tools are in use. Some of them are general purpose like Wikis, blogs, databases etc., and some are specific such as Moodle, Blackboard, CK-12, etc. All these tools provide features for E-learning explained by Osimo, D in [3]. The user activities are stored in databases and can be used for providing feedback to the overall learning process for further improvement.

E-learning is the key and primary term for all the educational learning associated with Web-based education explained by Cheng, S.-M. and Sue, P.-J., D in [4]. A system for E-learning is typically comprised of enormous amounts of knowledge and data that can be employed for scrutinizing students‟ behavior. It can be manipulated as a source of useful educational data, keeping track of all the students‟ actions including reading, writing, tests, peer communication and other related tasks. E-learning tracks the users‟ details which include all the interaction history and results within a secure database. Such databases are very difficult to maintain due to rapid data transition and generation explained by Porter, W. W. *et.al* in [5]. Although it has various beneficiary aspects, instructors would face certain difficulties in utilizing reporting tools. A few E-learning systems provide such tools, but it is very difficult for the instructor to gain useful information when it comes to a large number of students. E-learning systems could also lack in providing the access of students‟ activities to teachers. This condition affects the structure of the course and the learning process results are in paper by Mohamad, S. K. and Tasir, Z. [6]. Beside this, it is very useful for students to attain help from E-learning in terms of automatic guidance and recommendations in respect to on-line activities. Also some language issues pertain in use of E-learning materials. The E-learning material which is good may be present in language not understandable by the user.

Language translation in E-learning materials would not only favor better students‟ learning outcomes but also enable the learners to achieve specific goals online explained by Santos, O. C *et al* in [7]. One may categorize E-learning systems into three different categories: particular web-based courses, learning content management systems (LCMS) and adaptive and intelligent web-based educational systems.

1.1 Web-Based Courses

Particular web-based courses would specifically deal with the usage of standard HTML. This usage is intended to establish how the student uses the courses and how pedagogical strategies impact different types of students. These strategies involve, e.g. how many pages students have skipped, in which order they studied sub-topics, average time a student has spent on a single page and on the whole chapter, etc explained by Blagojević, M. and Micić, Z. in [8].

1.2 Learning Content Management Systems

Learning content management systems are platforms that offer a great variety of channels and workspaces to facilitate information sharing and communication between

participants in a course. Such systems let the instructor distribute information to students, produce content material, prepare assignments and tests, engage in discussions, manage distance classes and enable collaborative E-learning with forums, chats, file storage areas, news services, etc.

1.3 Adaptive and Intelligent Web-Based Educational Systems

The adaptive and intelligent web-based educational systems set the knowledge, structure and preferences on the basis of students interacting and evaluating their own needs. It is a mixed approach combining those of Intelligent Tutoring Systems (ITS) and Adaptive Hypermedia Systems (AHS) which is very efficient in terms of highlighting personalized learning experiences, student activities, providing feedback and the progress of students explained by Aleven V. in [9]. Two similar but different fields emerged: Educational Data Mining taken from IEDM, International Educational Data Mining Society cited in [10] and Learning Analytics from Society for Learning Analytics Research cited in [11].

2. The Benefits of Using E-Learning as a Support for Classroom Teaching

2.1 E-learning as a support for classroom teaching

ICT as a support for ordinary class-room teaching, and as a part of it, has the obvious benefits of: easy access whenever and wherever you wish it dematerialization (less paper – more trees) enabling us to use modern methodologies individualization (different interests/ levels/ needs).

2.2 The ethical dimension: learning to have a say

A Chinese proverb says: "teachers open the door, but you must enter by yourself." Our task is to encourage students. But it is not only the new vocabulary a good language course should give them. It is essential to consider the fact that the immediacy of the information and news reaching our students (through this new language) gives them an opportunity to be informed of and shape their opinion on important topics relating to our society and the "here and now".

As many renowned sociologist and researchers have stressed, it is not enough to have an opinion - an educated person must express it to shape the society we live in explained by Cronin M. in [12] and Pym, A. in [13]. Indeed, much of the knowledge and ideas in the modern 2.0/ 3.0 Web world are related to who has the information and who has it first. And who else should be encouraged to learn to use it to the best of their capacity and following all the ethical principles than students of foreign languages, who in many ways are and become the window to/ from the world of their own society and culture.

2.3 Building trust

Anatole France has said: "nine tenths of education is encouragement". There can be no encouragement without trust explained by [14] Kiggins J. & Cambourne B. in []. As teaching in general, so can also e-learning be organised in different ways. For some, it may be:

A ready-made environment created by the teacher teacher controlled, students present filled-in exercises, get marks self- tests "the teacher´s button" with which to control the computers of students when we work together in the computer classroom.

Much depends on the teacher´s authority type – whether the teacher is autonomy supporting or controlling. Dörnyei [15] points out: Sharing responsibility with students, offering them options and choices, letting them have a say in establishing priorities, and involving them in the decision-making process enhance student self-determination and intrinsic motivation.

2.4 Creating the feeling of belonging together

In education, as elsewhere, increased cooperation and neglecting of the earlier rigid borderlines, is becoming more and more of a common practice. Such an approach also helps students to retain their motivation. Cocea and Weibelzahl [16] point to the connection between e-learning and the Social Cognitive Learning Theory SCT. In their view, personalization, adaptivity, affective tutoring and collaborative learning, as well as motivation – all aspects:

Personalization aims to make learning more effective and satisfying by adapting to the learner‟s needs and preferences. Among the benefits of adapting to the learner‟s motivation are: enhanced motivation and involvement, empowered learners – making them more responsible and active, increased satisfaction, better quality of learning etc.

Motivation is related to affective computing, because self-concepts are always charged with emotions. Thus, affective agents could be used for both assessing motivation and intervention.

SCT also fits with collaborative learning, given the social framework taken in consideration by this theory and the way learning is influenced by the social context.

Rather contrary to what is sometimes supposed of a web-based environment, the experience shows that it often joins the students in the group. Offering them the possibility to communicate in an environment "natural" for them, the web- based course, if built up in a way that enables the students to participate and open up. It also serves to join the different terms (over X-mas, during the summer vacation) different parallel groups (e.g. Group A and Group B learning the same subject) different years of students learning the

same subject. This can be supported through helping students create common databases. In our case, the different power-point presentations and on-line dictionaries created by students have been the most popular items, especially so when the data-base is built up over different study years and together with the parallel group(s). Needless to say, the profiles of friends from a parallel group, and their small roster presentations of themselves also deserve great interest by the fellow students. And while commonly created on-line dictionaries can prove motivating for learning (and creating!) terminology and ESP for advanced students, consider how useful even creating a small roster presentation, or reading those of others, can be for a beginner-level general language student in the first months of their learning practice/ studies.

3. Proposed Work

The ICT use English as the medium for teaching the classroom students. The knowledge in books is rich and worth sharing. There are different books written in different regional languages. Also the old age knowledge used in mathematics and science use old Sanskrit language for its basic study. The translation of Sanskrit to native language takes a lot of human effort and time. The proposed system will directly translate the ICT used by the teacher and will present it to the students as shown in Fig. 1. Also there are different types of MT systems and approaches that can be used in ICT, explained in section IV and V.

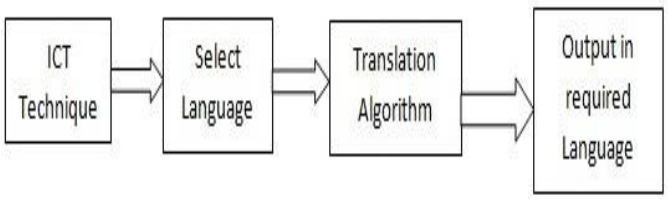

Figure 1. Proposed System

The translation algorithm can be implemented for translation the ICT through the databse. The sample algorithm is drawn in flowchart in figure 2 written by [17] Ruchika A. Sinhal and Kapil O. Gupta in [17].

The sentences from ICT documents may be feed to the training element. From the training element the system will adapt itself and will be ready for translating the sentences. The input from the ICT will be then be mapped to its respective translation and will be understood by the students and the trainer. [17]

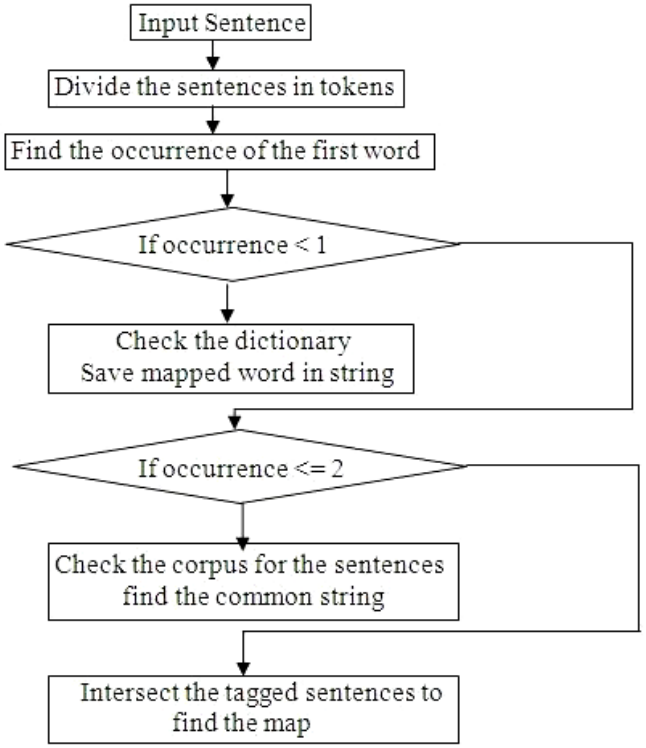

Figure 2. Flow chart for translation algorithm

4. TYPES OF MT SYSTEMS

For improving the E-learning translation can help students to learn. Smart E-learning can be implemented and performed by applying translation in ICT. The following are four types of Machine Translation (MT) systems explianed by Ruchika A. Sinhal and Kapil O. Gupta in [18]:

4.1 MT for Watcher (MT-W)

MT for watchers is intended for readers who wanted to gain access to some information written in foreign language who are also prepared to accept possible bad „rough" translation rather than nothing. This has been the type of MT envisaged by the pioneers. This came in with the need to translate military technological documents [19].

4.2 MT for revisers (MT-R)

MT for revisers aims at producing raw translation automatically with a quality comparable to that of the first drafts produced by human. The translation output can be considered only as brush-up so that the professional translator can be freed from that boring and time consuming task [19].

4.3 MT for translators (MT-T)

MT for translator"s aims at helping human translators do their job b y pro viding on-line dictionaries, thesaurus and translation memory. This type of machine translation system

is usually incorporated into the translation work stations and the PC based translation tools [19].

4.4 MT for Authors (MT-A)

MT for authors aims at authors wanting to have their texts translated into one or several languages and accepting to write under control of the system or to help the system disambiguate the utterance so that satisfactory translation can be obtained without any revision [19].

5. MT APPROACHES

Machine Translation is an attempt to automate, all or part of the process of translating one human language to another. The translation requires some knowledge of source and target languages and its way of interpretation to carry out the translation work. The MT systems can broadly be categorized on the basis of knowledge type, representation and interpretation of translation tools explained by Ruchika A. Sinhal and Kapil O. Gupta in [20].

The categories of MT systems are described in the next three sections. Since our research focuses on EBMT, this model is described in more detail by Sergei Nirenburg and Yorick Wilks in [21].

5.1 Knowledge Based MT

"The term knowledge based MT describe a system, displaying extensive semantic and pragmatic knowledge of domain, including an ability to reason to some limited extent, about concepts in the domain."

The basic aim of KBMT is to obtain high quality output in a specific domain with no post-editing work. The KBMT systems are generally domain specific, especially a domain that is less ambiguous, like technical documents. The reason for KBMT to be domain specific is that representing complete knowledge of the whole world is very difficult. The domain model is used to represent the meaning of the source language text.

The basic components of a KBMT system are:-

1. Ontology of the domain, which serves as an intermediate representation during translation. Ontology usually includes the set of distinct objects resulting from an analysis of a domain.

2. Source language lexicon and grammar for the analysis.

3. Target language lexicon and grammar for the generation.

4. The mapping rules between the intermediate and source/target language.

For example, the KANT system developed by CMT at Carnegie Mellon University is a practical translation system for technical documentation from English to Japanese, French and German by Hutchins, J. in [22].

5.2 Statistical MT

The researchers in the field of speech recognition first outlined the idea of statistical approach in machine translation. SMT is based on statistics derived from corpora of naturally occurring language, not with pre-fabricated examples. The view of the statistical approach is that every sentence in one language is a possible translation of any sentence of other language. The statistical model tries to find the sentence S in the source language for which the machine translator has produced a sentence T in the target language. This is based on the Bayesian or Noisy channel model used in speech recognition.

The model works with the intuition that the translated sentence has been passed through a noisy channel, which distorted the source sentence to the translated sentence. To recover the original source sentence we need to calculate the following –

1. The probability of getting the original sentence S in the source language.

2. The probability of getting the translated sentence T in the target language.

These are known as Language model and Translation model respectively. We assign to every pair of sentence (S, T) a joint probability, which is the product of the probability Pr(S) computed by the language model and the conditional probability Pr(T/S) computed by translation model. We choose that sentence in the source language for which the probability Pr(S/T) is maximum. Using Bayes theorem, we can write

where S = Source Text, T = Target Text, Pr (S/T) = probability that the decoder will produce S when presented with T, Pr(S) = probability that S would be produced in the source language, Pr(T/S) = probability that the translator will produce T when presented with S, and Pr(T) = probability that T would be Target language, but, here Pr(T) does not change for each S as we are looking for most-likely S for the same translation T.

In order to get the most-likely translation, we need to maximize Pr(S)*P(T/S). Thus, the formula to find the most likely translation T for a given sentence S is as follows –

The statistical system computes the language model probabilities (the probability of a word given, all the words preceding it in a sentence), the translation probabilities (the probability of the translation being produced) and uses a search method to find the greatest value (agrmax) for the product of these two probabilities thus giving the most probable translation.

5.3 Example Based MT

EBMT is a corpus based machine translation, which requires parallel-aligned three machine-readable corpora. Here, the already translated example serves as knowledge to the system. This approach derives the information from the corpora for analysis, transfer and generation of translation. These systems take the source text and find the most analogous examples from the source examples in the corpora. The next step is to retrieve corresponding translations. And the final step is to recombine the retrieved translations into the final translation. EBMT is best suited for sub-language phenomena like – phrasal verbs; weather forecasting, technical manuals, air travel queries, appointment scheduling, etc. Since, building a generalized corpus is a difficult task, the translation work requires annotated corpus, and annotating the corpus in general is a very complicated task. Nagao (1984) has been the first to introduce the idea of translation by analogy and claimed that the linguistic data are more reliable than linguistic theories by Makoto Nagao in [23]. In EBMT, instead of using explicit mapping rules for translating sentences from one language to another, the translation process is basically a procedure for matching the input sentence against the stored translated examples demonstrated by Ruchika A. Sinhal and Kapil O. Gupta in [20].

The basic tasks of an EBMT system are – - Building Parallel Corpora - Matching and Retrieval - Adaptation and Recombination The knowledge base, parallel aligned corpora consist of two sections, one for the source language examples and the other for the target language examples. Each example in the source section has one to one mapping in the target language section. The corpus may be annotated in accordance with the domain. The annotation may be semantic (like name, place and organization) or syntactic (like noun, verb, preposition) or both. For example, in the case of phrasal verb as the sub-language the annotations could be subject, object, preposition and indirect object governed by the preposition.

In the matching and retrieving phase, the input text is parsed into segments of certain granularity. Each segment of the input text is matched with the segments from the source section of the corpora at the same level of granularity. The matching process may be syntactic or semantic level or both, depending upon the domain. On syntactic level, matching can be done by the structural matching of the phrase or the sentence. In semantic matching, the semantic distance is found out between the phrases and the words. The semantic distance can be calculated by using a hierarchy of terms and concepts, as in WordNet. The corresponding translated segments of the target language are retrieved from the second section of the corpora. In the final phase of translation, the retrieved target segments are adapted and recombined to obtain the translation. The final phase identifies the discrepancy between the retrieved target segments with the input sentences‟ tense, voice, gender, etc. The divergence is removed from the retrieved segments by adapting the segments according to the input sentence‟s features. Let us consider the following sentences – - [Input sentence] John brought a watch. - [Retrieved - English] He is buying a book. - [Retrieved - Hindi] vaHa eka kitaba kharida raha he The aligned chunks are – [He] → [vaha] - [is buying] → [kharida raha he] - [a] → [eka] - [book] → [kitaba] The adapted chunks are – - [vaha] → [jana] - [kharida raha he] → [kharida] - [kitaba] → [gaghi] The adapted segments are recombined according to sentence structure of the source and target language. For example, in the case of English to Hindi, structural transfer can be done on the basis of Subject-Verb-Object to Subject-Object-Verb rule.

6. CONCLUSION AND FUTURE SCOPE

During our literature survey we reached the conclusion that E-learning is a modifiable, targeted, and highly approachable technique in educational systems, which leads to efficient and reliable learning and teaching results when used intelligently. This paper addresses targeted machine translation types and its approaches and their applicability in web-based E-learning. It provides an overview of recent studies in machine translation. Machine translation in E-learning no doubt can be implemented to make communication stable between instructor and students. Statistical data analysis provides a platform of reliable prediction about students‟ learning. From Web education-based courses to highly intelligent student prediction, it is proving itself as applicable to all operations. Machine translation in E-learning has scope for greater change and development. For instance, we have to rigorously work for the ethics preservation of students and teachers. Many techniques are to be developed to help foster students gain knowledge without having to work hard in learning other languages.

REFERENCES

[1] Clark, R. C. & Mayer, R. E., “e-learning and the Science of Instruction: Proven Guidelines forConsumers and Designers of Multimedia Learning,” 3rd ed. Pfeiffer, San Francisco, 2011.

[2] Greenhow, C., Robelia, B., & Hughes, J., “Learning, teaching, and scholarship in a digital age: Web 2.0 and classroom research: What path should we take now?,” Educational Researcher, 38, pp. 246– 259, 2009.

[3] Osimo, D., “Web 2.0 in Government: Why and How?,” JRC Scientific and Technical Reports, EUR 23358 EN, 2008.

[4] Cheng, S.-M. and Sue, P.-J. “Constructing concept maps for adaptive learning systems based on data mining techniques”, Expert Systems with Applications, Vol. 40, No. 7, pp. 2746-2755, 2013.

[5] Porter, W. W., Graham, C. R., Spring, K. A., and Welch, K. R. "Blended learning in higher education: institutional adoption and implementation", Computers & Education, Vol. 75, pp. 185-195, 2014.

[6] Mohamad, S. K. and Tasir, Z. "Educational data mining: a review", Procedia – Social and Behavioral Sciences, Vol. 97, No. 6, pp. 320-324 2013.

[7] Santos, O. C., Boticario, J. G., and Pérez-Marín, D. "Extending web-based educational systems with personalised support through User Centred Designed recommendations along the E-learning life cycle", Science of Computer Programming ,2013.

[8] Blagojević, M. and Micić, Z. "A web-based intelligent report E-learning system using data mining techniques", Computers & Electrical Engineering, Vol. 39, No. 2, pp. 465-474, 2013.

[9] Aleven, V. "Help seeking and Intelligent Tutoring Systems: theoretical perspectives and a step towards theoretical integration", in International Handbook of Metacognition and Learning Technologies, pp. 311-335 2013.

[10] IEDM, International Educational Data Mining Society, Available at http://www.educational datamining.org.

[11] Society for Learning Analytics Research,. http://www.solaresearch.org/"http://www.solare search.org.

[12] Cronin, M., "The Empire Talks Back: Orality, Heteronomy and the Cultural Turn in Interpreting Studies" , In F. Pöchhacker, M. Shlesinger (Eds.), The Interpreting Studies Reader. London and New York: Routledge. 393-397, 2002.

[13] Pym, A., "Action Research in Translation Studies" New Research in Translation and Interpreting Studies, International Studies Group, Tarragona, Spain, 7.-8. Oct. 2005.

[14] Kiggins, J. & Cambourne, B., "The knowledge building community program." In: T. Townsend, R. Bates (Eds.), Handbook of Teacher Education. Globalization, Standards and Professionalism in Times of Change. Springer, printed in the Netherlands. 365-381, 2007.

[15] Dörnyei, Z., Teaching and Researching Motivation. Pearson Education Limited. Malaysia, LSP, 2001

[16] Cocea, M. & Weibelzahl, S. (2006). Motivation – included or excluded from e-learning. http://www.easy-hub.org/stephan/cocea-celda06.pdf. Accessed October 2009.

[17] Ruchika A. Sinhal, Kapil O. Gupta,"A Pure EBMT Approach for English to Hindi Sentence Translation System", IJMECS, vol.6, no.7, pp.1-8, 2014.DOI: 10.5815/ijmecs.2014.07.01

[18] Ruchika A. Sinhal and Kapil O. Gupta, "Language Processing for MT: Need, Problems and Approaches" International Journal of Engineering Research and General Science Volume 1, Issue 1, August 2013 ISSN 2091-2730

[19] "Machine Translation –A Rosetta stone for the 21th century?", http://www.ida.liu.se/~729G11/projekt/student papper10/mariahedblom.pdf

[20] Ruchika Sinhal and Kapil Gupta, "Machine Translation Approaches and Design Aspects" IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, Ver. II (Jan. 2014), PP 22-25

[21] Sergei Nirenburg and Yorick Wilks, Machine Translation

[22] Hutchins, J. 1986. Machine Translation: Past, Present, Future, Ellis Horwood/Wiley, Chichester/New York.

[23] Makoto Nagao, A Framework of A Mechanical Translation between Japanese and English by Analogy Principle, In Artificial and Human Intelligence 1984, http://www.mt-archive.info/Nagao-1984.pdf