International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2020): 7.803

A Review on Intelligent Language Tutoring Systems& Relevance of NLP in ILTS

Lakshmi Kurup[1], Meera Narvekar[2]

Abstract: ***An Intelligent Language Tutoring System provides automated tutoring, pedagogical, feedback and assessment strategies for language learning. Language learning revolves around various facets like words, sentences, grammar, pronunciation, syntax and semantics. Mastering these language components requires a well-defined and formalized domain. There is a greater need to validate and structure the rules pertaining to syntactical, grammatical and semantical properties. However with the advancement of Artificial Intelligence and Natural Language Processing techniques, language tutoring systems can facilitate not only grammar and vocabulary based learning, but can also present students with auto generated questions and improvised learner-tutor interactions via natural language dialogue generation. The paper summarizes and briefly discusses the design of various ITSs for language learning and also attempts to analyze various Natural language Processing Techniques used in the design of ILTSs.***

Keywords: ITS, ILTS, NLP, LSA, RNN

1. Introduction

The research says the most effective way of teaching is one to one teaching wherein the computer tutors are found as good as human tutors. However, the research says that the active engagement of the student with an e-tutoring module is rewarding than the classroom teachingunder certain circumstances[1]. For language learning, a direct reference of dictionaries and vocabularies is not just enough. In language, as compared to other domains like mathematics, knowledge about a word and its context is not clearly defined. Also the presence of unseen or unknown vocabulary may affect the capacity of ITSs to interpret text in the user input. NLP Techniques focusses on processing and assessing the text input, which in turn makes it easier to diagnose the error and provide appropriate feedback. ILTSs attempt to develop NLP techniques that focus on processing and assessing the learner's free text input. With the advent of deep learning architectures like RNNs, the vector representations depicting the relations between the words like One-Hot vector representation, GLoVe representations, Continuous Bag of Words etc. has been extensively used in the area of language modeling. This in turn has paved way for developing better tutoring and pedagogical strategies in ILTS.

The ITSs have a conventional architecture consisting of four modules: Domain Module, which consists of facts and domain knowledge in the form of repositories or vocabulary lists; Student Module, which keeps track of student interactions, behaviors, responses, leaning patterns etc.; Tutor or a Pedagogical module, which focuses on teaching and enhancing tutor-student interactions using adaptive feedback and hint generations and finally, User Interface module which acts as an interface between the student and the tutor. In this paper we make an attempt to review various AI techniques used in each of these modules, their characteristics and evaluation methods.

The rest of the paper is structured as follows: Section II provides a review analysis on various Intelligent Language Tutoring Systems with focus on its modeling concepts.Section III explores various AI based language processing techniques used by some of the ILTSs. The paper is further concluded in Section IV.

2. A Review on ILTS

Significant progress has been made in the ITS field for well-defined domains such as algebra, physics, and computer programming. As far as English or any other languages are concerned languages are considered, because of the complexity and richness that allows speakers/writers to express themselves in so many ways, a tutoring system to teach the language has also got inherent difficulties. To perform knowledge acquisition and cognitive analysis, deep understanding of the domain knowledge, student behavior, student knowledge skills etc. are needed. Assessment and feedback generation is also a tedious process if the domain is ill defined. Table 1 reviews some of the ILTS dating from 1990's till now, along with its purpose and concepts used in design of domain model, student model and learning model.

Table 1: Analysis of Various ILTS.

S. No	ILTS	Purpose	Algorithms/Concepts on which domain/student/learning modeling is done
1	English Tutor[1] (1992)	Aims at learning of the English verb system. The proposed approach is based on the idea of analyzing the reasoning process of the student by in reverse order.	Domain Module consists of following tasks: parsing the exercise sentences, recognizing the temporal relations among the events described in the sentence, identifying the reference time for every clause in the sentence, selecting the correct tense to be used for each verb, conjugating the verb(s) into the appropriate tense. Student modeling is done by backward model tracing approach where each student response is matched with a set of predefined rules and if there is a mismatch, rules are generated dynamically based on previous examples.
2	CAPIT (2000)[2]	CAPIT is a Capitalisation And	CAPIT is based on Constraint-Based Modeling (CBM).The system

		Punctuation Intelligent Tutor designed for the 10-11 year old age group. CAPIT teaches a subset of the basic rules of English capitalisation and punctuation. .	represents the domain as a set of constraints specifying the correct patterns of punctuation and capitalisation, and feedback is given on violated constraints. A student model contains information about the history of previous sessions and also contains a model of the student's knowledge, expressed in terms of constraints.
3	Passive Voice Tutor [3] (2000)	The system may be used to present theoretical issues about the passive voice and to provide exercises that the student may solve	The domain knowledge performs following tasks: Parsing the exercise, Identifying the semantic relationships between the subject, the verb and the object and judging whether the exercise sentence makes sense. Even if the sentence is semantically invalid, the system can convert to passive voice, if asked to.
4	REAP[4](2004)	The system is used to enhance the reading skills of students.	Domain knowledge is constructed by an advanced search model which finds documents that satisfy complex lexical constraints, vocabulary unknown to the student etc. Data and Linguistic annotations are done offline at indexing time. POS tagging, named entity recognition and sentence level parsing is done on the indexed documents.
5	TAGARELA, Teaching Aid for Grammatical Awareness, Recognition and Enhancement of Linguistic Abilities[5]	TAGARELA can be viewed as an intelligent electronic workbook. Its exercise types are similar to the ones found inregular workbooks:It has exercised on -Listening -Reading -Description -Fill in the Blanks -Rephrasing -Vocabulary	As it receives input from the user, it decides on which NLP modules to call. It uses WordNet modules for tokenization, spell checking and shallow parsing
6	ICICLE, Interactive Computer Identification and Correction of Language Errors [6](2006)	The system is designed for deaf learners to track their proficiency in their English writing skills. .	The analysis process relies on lexical information from the COMLEX Syntax 2.2 lexicon, the Kimmo morphological processor, Allen's TRAINS parser version 4.0 with an augmented CFG grammar which has been developed specifically for ICICLE
7	BEETLE II, Basic Electricity and Electronics Tutorial Learning Environment[6] (2014)	The system understands the student responses and provides context-specific feedback based on Naturallanguage understanding.	The system uses a standard interpretation pipeline and domain-independent parsing for feedback generation. A deep parser is used to analyze the student responses. A contextual interpreter is uses to convert parser output and then sent to tutorial planner, which in turn decides which tutorial tactics to use and finally the NLG component converts the tutorial tactics to textual representation and presents it to user.
8	An Intelligent Tutoring System for Teaching English Grammar [7](2016)	The system aims to tach basics of English Grammar.	The domain model contains information relating to the English grammar tenses or number of strategies to represent information. The student model forms the frame which determines the current state of the student in understanding English grammar tenses. The student knows status through the creation of profile.

The ILTSs mentioned in the table are more constrained to NLP and its APPLICATIONS. There are quite a few more ILTSs listed when NLP techniques are elaborated in detail.

3. Use of NLP in ILTS

The use of NLP in ILTS dates back to 1960s.Different ITSs uses various natural language processing techniques based on its application.NLP is used to automate many tasks like analyzing the structure, verification of errors, generation of feedback messages, content creation as per curriculum demand, generation of multiple choice questions etc.But there are many challenges with NLP too. First, because many NLP tools have been trained on professionally written texts such as the Wall Street Journal, they often do not perform well when applied to texts written by students. Second, we cannot have a generalized variable, nut a more pedagogically relevant variable. For example, although word count can be used to generate essay scores, word count is typically not part of a human's grading rubric and would thus not be useful to mention in student feedback. Some of the techniques used w.r.t to ILTS are mentioned below.

Syntactic Analysis:
Syntactic analysis can be used to find out the errors in the writing, prepositional errors etc. ICICLE[8] does syntactic analysis to improvise the writing skills of the deaf students. Standard proof reading tools cannot be relied to detect these type of syntactical errors.

Semantic Analysis:
Semantic Analysis is incorporated when we need to assess the similarity and meaning between the student responses and reference answers, Paraphrase assessment, essay grading etc.Intelligent tutors fails to evaluate semantic knowledge as they assume that learners acquire the semantic knowledge before doing the exercises.But if the learner possess wrong semantic knowledge then the tutoring module strategies would turn inappropriate. Semantic knowledge is generally stored or achieved by asking direct

questions about that knowledge as in multiple choice questions or true/false questions. Semantic KB can also be updated by concept maps that a learner would draw where nodes represent a concept and links represents relationship between the concepts [9]. Weighted bag of words, first order logic with unification, syntactic graph distance are some of the entailment techniques used to make semantic inferences.Some systems are based on grammatical rules and logic whereas some uses statistical techniques like Word matching, Latent semantic Analysis (LSA),[10].

a) Word Matching

CIRCSIM-Tutor V2 is a dialogue based ITS, which uses natural language for input and output and can perform syntactic check on sentence fragments and identify misspelled words. It uses a word matching algorithm on both the unknown input and the candidate word from the lexicon. It also generates text from logic forms which has fill in the blanks templates for insertion of auxiliary verbs, active/passive transformations and finally the sentence gets linearized using a Lexical Functional Grammar [11].

b) Latent Semantic Analysis

LSA is a statistical technique used for extracting the relations between words based on their common occurrences in texts. LSA stores each single text in a corpora in the form of a frequency matrix. It has a greater disadvantage that it lacks basic logic of language constructs like syntax, negation, word order etc. and so it does not yield any verbal intelligence in real world phenomena.[12]

LSA-based intelligent tutoring systems are majorly used to model conversational dialogues in AutoTutor(2004)[13]. It incorporates a curriculum script repository, where the content related to a question or problem is stored. Each entry includes an ideal answer, sets of expectations and misconceptions, hints and prompts, a set of key words and synonyms, and a summary.

LSA has its significance in enhancing text comprehension and summarization skills in State-The-Essence Summary Street (2001) [14] ITS .It also has been applied to improvise story and essay writing in Select-a-Kibitzer(2000)[15], Story Station (2003) [16] ITS.

iSTART (Interactive Strategy Trainer for Active Reading and Thinking) is a web-based, automated tutor designed to help students become better readers using multi-media technology[17]. LSA opened up a lot of possibilities for above mentioned tutoring systems such that they were able to automate natural language and generate appropriate feedback statements.

In recent years dialogue-based intelligent tutoring systems have become more prevalent as it enhances learning in one-one tutoring. Many well defined domains like biology, physics, circuit design, thermodynamics uses natural Language generation to provide user responses. NLP takes advantage of increasing amount of digital data in the form of text and speech from blogs, social media, websites, logs etc for content creation and auto generation of multiple choice questions, word banks etc.

4. Conclusion

Intelligent Tutoring Systems (ITSs) plays a significant role in simulating the classroom teaching thereby enhancing personalized tutoring. Various ILTSs and the concepts used in modeling the domain model, student model and pedagogical models were discussed. NLP helps ILTS to evaluate and process the learner's input text and provide appropriate pedagogical feedback. Furthermore, current NLP methods are often tailored to data that are too context specific, restricted to a vocabulary set trained on professionally written text like WSJ, and well defined STEM domains like physics and mathematics.

References

[1] Tasso C., Fum D., Giangrandi P."The Use of Explanation-Based Learning for Modelling Student Behavior in Foreign Language Tutoring," in *Springer-Verlag Berlin Heidelberg*, 1992.

[2] Michael Mayo.,Antonija Mitrovic, Jane McKenzie, "CAPIT: an intelligent tutoring system for capitalisation and punctuation," in *Advanced Learning Technologies, 2000. IWALT 2000. Proceedings*, February 2000.

[3] Virvou Maria, Maras Dimitris, Tsiriga Victoria , "Student Modelling in an Intelligent Tutoring System for the Passive Voice of English Language," *JSTOR*, vol. Vol. 3, no. No. 4, pp. 139-150, October 2000.

[4] Kevyn Collins-Thompson ,J.Callan, "Information Retrieval for Language Tutoring:An Overview of the REAP Project," in *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, July 2004.

[5] "Feedback Grammatical Systems for Introduction Language Tutoring," *European Association for Computer-Assisted Language,EUROCALL*, no. Issue number 10, September 2006.

[6] Myroslava Dzikovska, Natalie Steinhauser, Elaine Farrow, Johanna Moore & Gwendolyn Campbell , "BEETLE II: Deep Natural Language Understanding and Automatic Feedback Generation for Intelligent Tutoring in Basic Electricity and Electronics," *International Journal of Artificial Intelligence in Education*, p. 284–332, 10 June 2014.

[7] Mohammed Ibraheem Alhabbash, Ali Osama Mahdi, Samy S. Abu-Naser, "An Intelligent Tutoring System for Teaching Grammar English Tenses," *European Academic Research* , vol. Vol. IV, no. Issue 9, pp. 1-6, December 2016.

[8] Lisa N. Michaud, Kathleen F Mccoy , "Capturing the Evolution of Grammatical Knowledge in a CALL System for Deaf Learners of English," *International Journal of Artificial Intelligence in Education, IJAIED*, vol. 16, no. 1, pp. 65-97, 1 Nov 2005.

[9] Philippe Fournier-Viger, Roger Nkambou and André Mayers, "A Framework for Evaluating Semantic Knowledge in Problem-Solving-Based Intelligent Tutoring Systems.," in *Conference: Proceedings of the Twenty-First International Florida Artificial Intelligence Research Society Conference*, Coconut Grove, Florida, USA, May 15-17, 2008.

[10] Chutima Boonthum-Denecke, Irwin B. Levinstein, Danielle S. McNamara, Joseph P. Magliano and Keith K. Millis, Hampton University, NLP Techniques in Intelligent Tutoring Systems, USA: Encyclopedia of Artificial Intelligence, 2009, p. 6.
[11] Nakhoon KimMartha EvensJoel A. MichaelAllen A. Rovick, "CIRCSIM-Tutor: An Intelligent Tutoring System," in *Computer-Assisted Learning: 2nd International Conference*, Berlin, 1989.
[12] W. Lenhard, "Bridging the Gap to Natural Language: A Review on Intelligent Tutoring Systems based on Latent Semantic Analysis," Research Report, Germany, 2008.
[13] Arthur C. Graesser, Shulan Lu, George Tanner Jackson, Heather Hite Mitchell, Mathew Ventura, Andrew Olney & Max M. Louwerse , "Using LSA in AutoTutor: Learning Through Mixed-Initiative Dialogue in Natural Language," *The Handbook of Latent Semantic Analysis (pp).*, pp. 243-262, 2007.
[14] David Wade-Stein, Eileen Kintsch , "Summary street: interactive computer support for writing," Cognition and Instruction, 22 (3) , (2004).
[15] PeterWiemer-Hastings ArthurC.Graesser , "Select-a-Kibitzer: A computer tool that gives meaningful feedback on student compositions," Special Issue of Interactive Learning Environments, Memphis TN, 2000.
[16] Judy Robertson and Peter Wiemer-Hastings , "StoryStation: an intelligent tutoring system for story writing," 2003.
[17] D. M. S. D. &. W. T. Landauer, A Handbook of Latent Semantic Analysis Mahwah,Journal of Language and Social Psychology, 27(3):300-306, NJ: Erlbaum, May 2008
[18] A. C. Graesser, K. VanLehn, C. P. Rose, P. Jordan and D. and Harter, Intelligent tutoring systems with conversational dialogue, 2001.
[19] L. N. M. Kathleen F McCoy, "Capturing the Evolution of Grammatical Knowledge in a CALL System for Deaf Learners of English," International Journal of Artificial Intelligence in education(IJAIED), Vols. Vol. 16,No:1, pp. 65-97, 2006.