A Cognitive Tutor of Arabic Word Root Extraction Using Artificial Word Generation, Scaffolding and Self-Explanation

Hanan Elazhary
King Abdulaziz University, Jeddah, Saudi Arabia;
Electronics Research Institute, Cairo, Egypt
helazhary@kau.edu.sa; hananelazhary@eri.sci.eg

Nabila Khodeir
Electronics Research Institute, Cairo, Egypt
nkhodeir@eri.sci.eg

Abstract—Extracting the roots (stemming) of Arabic words is one of the most challenging skills taught to Arabic language learners. To address this challenge, this paper proposes the Arabic word Root extraction Tutor (ART). ART is a cognitive tutor intended to teach students production rules needed for Arabic word root extraction. In the passive mode, ART accepts an input word and generates its root with explanation of the correct production rule. In the active mode, on the other hand, words are generated by ART and the student is prompted to provide the correct roots. ART integrates several techniques for enhanced tutoring. It provides a positive feedback for a correct answer and a negative one otherwise. In the latter case, Prompting Answer Strategy (PAS) is employed, where the student is guided to detect the error by integrating scaffolding and self-explanation. Scaffolding prompts the student to apply the relevant production rule step by step. In each step, a number of options are given to the student to select the correct one via self-explanation. If the error persists, the correct root is generated with explanation of the correct production rule. In addition to generating real words, artificial words are generated using the production rules. This novel technique is intended to ensure that the student applies the production rules rather than memorizes the roots of common words. Evaluation has shown the effectiveness of ART tutoring process and suggests artificial word generation as a promising technique in language tutoring.

Keywords—Arabic, Artificial Words, Cognitive Tutor, Intelligent Tutoring Systems, Problem Generation, Root Extraction, Scaffolding, Self-explanation, Word Stemming.

1   Introduction

Intelligent Tutoring Systems (ITS) aim at tutoring students in the absence of or in addition to human tutoring. A typical ITS generates and/or accepts problems relevant

to a given domain and attempts to analyze the student error and build a short-term student model to provide relevant feedback and help. A more advanced ITS builds a long-term student model to provide the student with personalized problems according to the assessed level of knowledge. Several approaches to short-term and long-term student modeling have been proposed over the years [1]. These include: generative modeling [2], overlay modeling [3], enumerative bug modeling [4], constraint-based modeling [5] and cognitive model tracing [1, 6-7].

Cognitive tutors [1, 6-7] are based on the *ACT-R* theory of cognition that assumes that human beings own declarative and procedural long-term memory stores. The declarative memory store is used to store chunks of declarative and factual knowledge. Learning converts this knowledge into procedural knowledge stored in the procedural memory store. The procedural knowledge is represented by production rules needed to achieve corresponding goals. As the learner becomes an expert, the production rules are optimized. Consequently, cognitive tutors are suitable for domains in which human experts' knowledge and how they solve problems can be represented using a set of If-Then production rules. They are intended to tutor the students how to exploit the production rules for solving problems in a given domain according to the goals. Production rules have the advantage that they can be used by the automated tutor to generate the correct answer of a given problem in the same way the expert does and consequently, they allow tracking the problem solving steps of the student. Short-term student modeling is achieved via model tracing, in which an error is detected whenever the student answer or answer step does not match any of the production rules.

The challenging skill of Arabic word root extraction (stemming) is a very good candidate for cognitive tutors since knowledge in this domain is essentially represented by a set of If-Then production rules explaining how the roots of the different words can be extracted according to their types and structures. This paper proposes the Arabic word Root extraction cognitive Tutor (ART), which works in passive and active modes. In the passive mode, the student can input a word and the system responds by generating the root and an explanation of the relevant production rule. This mode can be used by the student for self-paced learning. In the active mode, on the other hand, the system outputs a word and the student is prompted to provide its root. ART employs a novel technique in language learning. In addition to generating real words, artificial words are synthesized using the production rules to ensure that the student applies the production rules rather than memorizes the roots of common words.

ART employs several techniques to help in tutoring the student in the active mode. It generates a positive feedback in case of a correct answer and a gentle negative feedback otherwise. It has been shown that students who receive both positive and negative feedback learn two times faster than the students who receive negative feedback only [8]. In the latter case, ART does not provide the student with the correct answer directly. Alternatively, it employs Prompting Answer Strategy (PAS) where the student is guided to recognize the error and provide the correct answer. It has been shown that PAS is effective in computer-assisted language learning [9]. In ART, PAS is realized through integrating scaffolding and self-explanation. Scaffolding [10] is achieved by guiding the student to apply the relevant production rule step by step. In

each step, the student is provided with several options to select the correct one using self-explanation. It has been shown that the students who employ self-explanation while learning learn much faster than those who do not [11]. If an error persists, the correct root is generated accompanied with an explanation of the corresponding production rule. The contributions of this paper can be summarized as follows:

- Developing the cognitive tutor, ART for Arabic word root extraction that can operate in both the active and passive modes.
- Integrating positive and negative feedback, PAS, scaffolding and self-explanation for effective tutoring of Arabic word root extraction.
- Proposing the novel technique of generating artificial words using the production rules to make sure the students are able to apply the rules even with unfamiliar words and demonstrating via empirical evaluation that this is a promising technique in language learning.

The paper is organized as follows: Section 2 provides related research in the literature in order to highlight the contributions of the paper. Section 3 explains the production rules employed by ART for Arabic word root extraction. Section 4 explains how ART generates real and artificial words. Section 5 discusses student model tracing in ART while Section 6 provides an overview of the system. Section 7 presents the results of the empirical evaluation of ART and the novel technique of artificial word generation. Finally, Section 8 provides the conclusion of the paper and directions for future research and enhancements.

2   Related Work

Several algorithms have been proposed in the literature for the automated extraction of the roots of Arabic words. For example, in the technique proposed by Al-Shalabi and Evens [12] a database of roots is pre-prepared. To extract the root of an input word, the word is first pre-processed morphologically to remove the longest possible prefix from it. According to the production rules of Arabic word root extraction, the roots exist in possible sets of locations in the remaining portion of the word. Those sets of locations are thus compared to the roots in the database to find a matching root. Abu Hawas [13] utilized the production rules in a different way. In her proposed algorithm, to extract the root of an input word, she tries to predict the word letters that comprise this root by comparing the relations among the word letters to the production rules. Yousef et al. [14] tried to avoid using a large number of production rules using N-gram. Kanaan and Kanaan [15] proposed an algorithm that examines the letters of an input word one by one starting with the last letter backwards. Those letters are matched against a set of production rules that exploit important morphological aspects of the Arabic language.

In spite of all this effort, none of those algorithms reported 100% accuracy. This was emphasized by Al-Shawakfa et al. [16] who carried out a comparative study of six recent or popular Arabic word root extraction algorithms. In this study, the highest reported accuracy was only 55%. This stems from the fact that most of the algorithms

developed in the literature are heuristic algorithms that favor speed over accuracy and simplicity and so they are not intended or suitable for tutoring. Elazhary et al. [17] tackled this problem by developing an automated tutor that utilizes a set of exact production rules for Arabic word root extraction. The set they adopted is concerned with past-tense Arabic verbs whose roots are of length three or four letters. The rules are simplifications of formal rules to be easily learnt, but are exact and are thus accurate. Nevertheless, their proposed tutor works in the passive mode only. In other words, it accepts an input word and generates its root and optionally the corresponding formal and simplified production rules.

Table 1. Simplified Rules for Extracting the Roots of Past-Tense Arabic Verbs Composed of Three or Four Letters (altered from [17])

Rule	Verb	Composition	Root
Rule 1: If the verb is composed of three letters and the first letter is آ, replace آ with أ	آكل	آ ك ل	أكل
Rule 2: If the verb is composed of four letters and the second and third letters are similar, delete one of the similar letters	علّم	ع ل ل م	علم
Rule 3: If the verb is composed of four letters and the second letter is ا, delete ا	قاتل	ق ا ت ل	قتل

This paper proposes the cognitive automated tutor, ART, which adopts the same set of production rules, but works in both the passive and active modes. It can generate both real and artificial words. It also utilizes PAS and integrates scaffolding and self-explanation for an enhanced tutoring process as explained in the following sections.

3 ART Production Rules

As previously noted, the current version of ART adopts a set of fifteen production rules needed for extracting the roots of Arabic past-tense verbs whose roots are of length three or four letters [17]. A motivation behind selecting this set of rules is that those rules are not typically taught to students as part of the curriculum. Alternatively, the students are typically taught a single rule stating that "*If the verb has extra letters, remove those letters to obtain the root.*" In spite of the simplicity and the correctness of this abstract rule, it does not do much help in teaching the students how to extract the root of a given verb since it is not clear how to identify the extra letters of that verb. Detailed formal rules do exist, such as the rule shown in Figure 3 (in Arabic), but they need to be simplified to be easily applicable.

The set of fifteen rules adopted by ART identify the extra letters by providing simplified explanation of which letters should be removed from a given verb depending on the number of its letters and its composition. Nevertheless, replacing such a single abstract rule with fifteen detailed rules (regardless of their simplicity) implies the need to provide the students with substantial training to be able to apply them effectively and efficiently. This makes this set of rules a perfect candidate for adoption by

ART and justifies the need for ART as a valuable tutoring tool for providing the required training.

Table 2. Simplified Rules for Extracting the Roots of Past-Tense Arabic Verbs Composed of Five Letters (altered from [17])

Rule	Verb	Composition	Root
Rule 4: If the verb is composed of five letters, the first letter is ا, and the fourth and the fifth letters are similar, delete ا and one of the similar letters	احمرّ	ا ح م ر ر	حمر
Rule 5: If verb is composed of five letters, the first letter is ت and the third letter is ا, delete ت and ا	تعامل	ت ع ا م ل	عمل
Rule 6: If the verb is composed of five letters, the first letter is ت and the third and fourth letters are similar, delete ت and one of the similar letters	تفهّم	ت ف ه ه م	فهم
Rule 7: If the verb is composed of five letters and the first letter is ت, delete ت	تدحرج	ت د ح ر ج	دحرج
Rule 8: If the verb is composed of five letters, the first letter is ا and the second letter is ن, delete ا and ن	انكسر	ا ن ك س ر	كسر
Rule 9: If the verb is composed of five letters, the first letter is ا and the third letter is ت, delete ا and ت	انتصر	ا ن ت ص ر	نصر

Tables 1, 2 and 3 provide the fifteen simplified production rules adopted by ART. Each rule is accompanied by an example verb, the letters it is composed of and its corresponding root for clarification of how to apply the rules. In the next sub-section, we discuss some properties of the Arabic language that help readers in following the examples in the tables.

3.1 Some Characteristics of the Arabic Language

It is worth noting that the words and sentences of the Arabic language are written and read from the right to the left. Besides, the Arabic letters generally take on different forms depending on their positions in the words and the identity of the neighboring letters. Additionally, the Arabic language exploits diacritics, which are marks above or below the letters, to make up for the absence of short vowels. The *shadda* diacritic ٌ indicates a repetition of the corresponding letter. In other words, it indicates the existence of two similar letters. This explains the following:

- The second letter of the verb علّم shown in Table 1 (Rule 2) is repeated as indicated by the *shadda* implying the presence of two similar letters.
- The second letter of the verb قاتل shown in Table 1 (Rule 3) is in fact the letter ا.
- The fourth letter of the verb احمرّ shown in Table 2 (Rule 4) is repeated.

- The first and the third letters of the verb تعامل shown in Table 2 (Rule 5) are the letters ت and ا respectively. Besides, its second letter is the letter ع, which takes on a different form when it appears as the first letter of the root عمل.
- The first letter of the verb تفهّم shown in Table 2 (Rule 6) is the letter ت and its third letter is repeated.
- The first letter of the verb تدحرج shown in Table 2 (Rule 7) is the letter ت.

Table 3. Simplified Rules for Extracting the Roots of Past-Tense Arabic Verbs Composed of Six Letters (altered from [17])

Rule	Verb	Composition	Root
Rule 10: If the verb is composed of six letters, the first letter is ا and the fourth and fifth letters are similar, delete ا and the similar letters	اجلوّذ	ا ج ل و و ذ	جلذ
Rule 11: If the verb is composed of six letters, the first letter is ا and the fourth letter is ن, delete both of ا and ن	افرنقع	ا ف ر ن ق ع	فرقع
Rule 12: If the verb is composed of six letters, the first letter is ا, the fourth letter is ا and the two last letters are similar, delete the first ا, the fourth ا and one of the similar letters	اصفارّ	ا ص ف ا ر ر	صفر
Rule 13: If the verb is composed of six letters, the first letter is ا and the two last letters are similar, delete ا and one of similar letters	اطمأنّ	ا ط م أ ن ن	طمأن
Rule 14: If the verb is composed of six letters, the first letter is ا, the fourth letter is و and the third and fifth letters are similar, delete ا, و and one of the similar letters	اخشوشن	ا خ ش و ش ن	خشن
Rule 15: If the verb is composed of six letters, the first letter is ا, the second letter is س and the third letter is ت, delete ا, س and ت	استخرج	ا س ت خ ر ج	خرج

- The second letter of the verb انكسر shown in Table 2 (Rule 8) is the letter ن.
- The third letter of the verb انتصر shown in Table 2 (Rule 9) is the letter ت.
- The fourth letter of the verb اجلوّذ shown in Table 3 (Rule 10) is repeated.
- The fourth letter of the verb افرنقع shown in Table 3 (Rule 11) is the letter ن.
- The fourth letter of the verb اصفارّ shown in Table 3 (Rule 12) is the letter ا and its last letter is repeated.
- The last letter of the verb اطمأنّ shown in Table 3 (Rule 13) is repeated.
- The second and the third letters of the verb استخرج shown in Table 3 (Rule 15) are the letters س and ت respectively.

It is clear that those rules are cumbersome. They require recognizing the number of letters of each verb and specifically, realizing that the *shadda* diacritic ◌ّ indicates a

repetition of the corresponding letter. They also require realization of fifteen different possible structures of a given verb and knowledge of how to apply the corresponding rules efficiently and effectively.


	(a) Rule 11	(b) Rule 4
Template		
Basic letters	ن ا	ا
All letters	ع ط م ن ش ا	ا ف ش ط ط
Generated word	اشمنطع	افشطّ
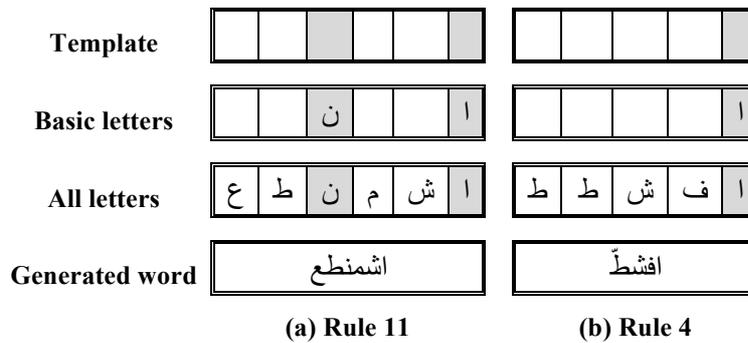
Fig. 1. Examples of ART Artificial Word Generation Using (a) Rule 11 & (b) Rule 4.

4 Word Generation in ART

As mentioned before, ART can generate both real and artificial words. Real words corresponding to each rule are stored in a data store and to generate a real word, ART selects words randomly from this data store. Unfortunately, no matter how large the size of the data store is, there is still a chance that the student memorizes the corresponding roots rather than apply the production rules to generate them. For this reason, ART generates artificial words whose roots are unknown to the student. To generate artificial words, ART exploits the production rules. To generate a word using a given production rule, ART does the following:

- First it creates a template with a number of positions equal to the number of letters mentioned in the rule
- Next, it inserts the basic letters mentioned in the rule in their corresponding positions in the template.
- It randomly selects distinct letters equal in number to the empty positions in the template and inserts each of them in its corresponding empty position.
- In case two consecutive letters should be similar, this is taken into consideration and the letters are replaced by a single letter modified by the *shadda* diacritic.

To illustrate, Figure 1 shows examples of the process of artificial word generation corresponding to Rule 4 and Rule 11. In each example, the figure shows the empty template, the template after inserting the basic letter(s), the word after inserting random letters and the final generated word. For example, Figure 1(b) shows the template corresponding to Rule 4 with five empty positions equal to the number of letters mentioned in the rule. As mentioned in the rule, the first letter of a corresponding word should be an ا. Accordingly, this basic letter ا is inserted in the first position of the template. Since the template has four empty positions, ART normally randomly

selects four distinct letters to fill in the empty positions. Nevertheless, according to Rule 4, the letters in the 4th and the 5th positions should be similar. Thus, a single letter is selected for both positions. In other words, three distinct letters are inserted in the 2nd, 3rd and 4th positions of the template (which are the first three empty positions of the template) and the letter in the 4th position is duplicated in the 5th position. Those two similar letters are then replaced by a single letter modified by the *shadda* diacritic resulting in the artificial word افشطّ corresponding to Rule 4.

5 Model Tracing in ART

This section explains model tracing in ART. Model tracing is employed to detect and diagnose the student error and provide relevant help in case the student answer doesn't follow the correct production rule. As mentioned before, in case of an error, PAS is employed to prompt the student to detect the error and provide the correct answer. PAS is realized through integrating scaffolding and self-explanation, where the student is guided to apply the relevant production rule step by step.

Fig. 2. ART main interface.

As shown in Tables 1, 2 and 3, the first step in applying any production rule is to count the number of word letters. This is pretty easy in most cases except in some cases such as when one of the letters is modified by a *shadda*. For example, suppose that the student is provided with the word احمرّ shown in Table 2 (Rule 4). Unless the student realizes that the *shadda* diacritic indicates a repetition of the corresponding letter ر, he/she will not be able to count five letters for the word and so will not be able to apply the correct production rule. Thus, in the first scaffolding step, the student is provided with a number of options about the number of letters in the word to select the correct one via self-explanation. A possible set of options is the following:

- The word is formed of four letters ا حمـر
- The word is formed of five letters ا حمـرر since the *shadda* indicates a repetition of the letter ر
- The word is formed of five letters ا حمـمـر since the *shadda* indicates a repetition of the letter م

If the student does not select the correct option, which is the second option in this example, ART deduces that the student is unaware of the role of the *shadda* and so generates an explanation to the student. The student is then prompted again to provide the correct root. If the error persists, the student is taken to the next scaffolding step.
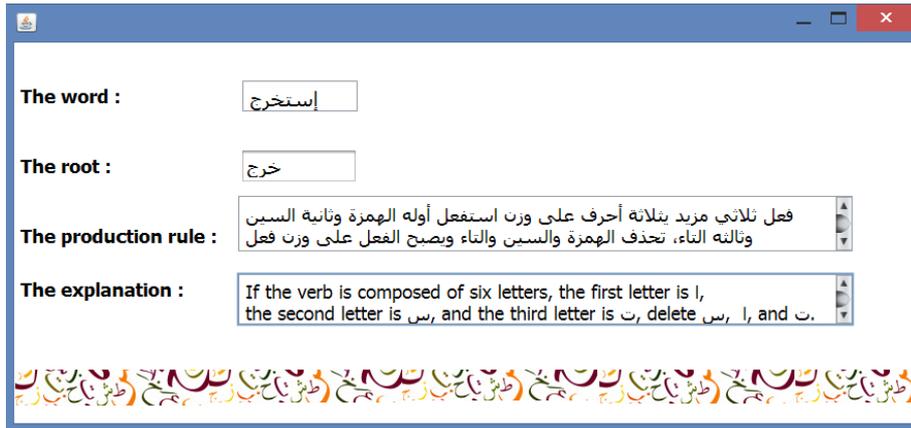
Fig. 3. A Snapshot of ART in The Passive Mode.

As shown in the tables, the second step in the application of the production rules involves realizing the structure of the word. Thus, the student is again provided by a set of corresponding options to select the correct one via self-explanation. A possible set of options is the following:

- the first letter is ا, and the fourth and the fifth letters are repeated
- the first letter is ا and the second letter is ح
- the first letter is ا and the third letter is م

Although the word structure is obvious and those options might seem to be redundant, they are still useful in guiding the student to realize the structure that corresponds to a production rule. Additionally, in a rare case, the student may not be able to realize a letter since, as mentioned before, the form of an Arabic letter may change with the change of its position in the words. This is more likely to occur with students learning Arabic as a second language. Again, if the student is unable to make the correct choice, which is the first option in this example, ART provides an explanation. The student is then prompted again to provide the correct root and if the error persists, the student is taken to the next scaffolding step.
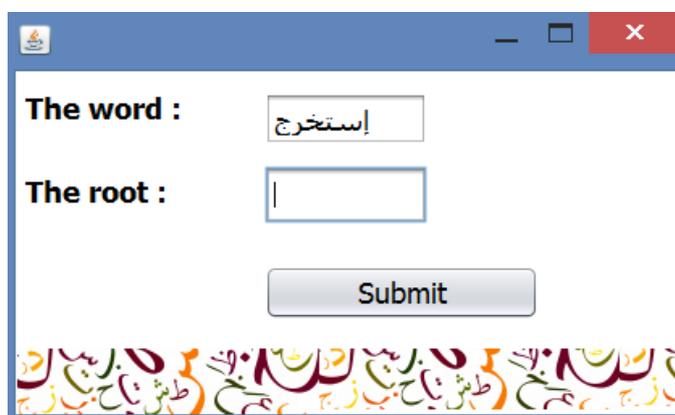
Fig. 4. A Snapshot of ART in The Active Mode

The next step in providing the correct answer is determining the correct production rule. Thus, the student is provided with a set of corresponding options to select the correct one. A possible set of options is the following:

- If the verb is composed of five letters, the first letter is ا and the second letter is ح, delete ا and ح
- If the verb is composed of five letters, the first letter is ا, and the fourth and the fifth letters are similar, delete ا and one of the similar letters
- If the verb is composed of five letters, the first letter is ا and the third letter is م, delete ا and م

Again, in case the student is unable to make the correct choice, which is the second option in this example, ART generates the relevant production rule. The student is then prompted again to provide the correct root. If the error persists, ART generates the correct root with explanation of the correct production rule.

6   Overview of ART

This section provides an overview of ART. Figure 2 shows the main interface of ART. As mentioned before, ART operates in two modes, passive and active. In the passive mode, the student clicks the button *Input a word* to input a word to the system. In the active mode, on the other hand, the student clicks the button *Learn* and the system outputs a word for testing the student knowledge.

Figure 3 shows a snapshot of ART in the passive mode. The student input the word استخرج and ART responded by generating the root, the formal production rule and the simplified production rule used in generating the root as an explanation.

Figure 4, on the other hand, shows a snapshot of ART in the active mode. As mentioned before, in this mode, the system outputs a word and the student is prompted to

provide its root as shown in the figure, where the word استخرج is output by the system. In case the student inputs the correct root, a positive feedback is provided to encourage the student to learn and improve.

On the other hand, in case the student inputs an incorrect root, ART generates a gentle negative feedback and employs PAS to guide the student to recognize the error and input the correct root rather than providing him/her with the correct answer directly. As mentioned before, in ART, PAS is realized through scaffolding and self-explanation. Scaffolding is achieved by guiding the student to apply the relevant production rule step by step. In each step, the student is provided with several options to select the correct one via self-explanation.

As explained in Section 5, scaffolding continues to guide the student to realize the structure of the word and to remember and apply the correct production rule to provide the correct root. The process ends when the student provides the correct answer. Nevertheless, if the error persists and the student provides an incorrect root, ART generates the correct root together with the formal production rule and the simplified rule used in generating the root similar to the passive mode.

7 Empirical Evaluation

To assess the effectiveness of ART model tracing and the novel technique of artificial word generation, two groups of students were formed. Each group was formed of twenty students in the age range of 15 to 22. The students were given a one-hour lecture explaining the production rules. Afterwards, they were given a pre-test exam and the grades were recorded. The students were then allowed to be trained using ART for two hours, but the artificial word generation was disabled in the system used by the second group. They were then given a post-test exam.

7.1 Evaluation of ART Model Tracing

Paired samples *t-test* was applied to compare the grades of the pre-test exam and the post-test exam of each group. For the first group, our hypothesis H_0 was that there is no difference between the grades of the exams pair. The test was performed using Real Statistics [18] and repeated using Social Science Statistics [19]. We obtained a *t-value* of 8.03. Using this *t-value* and degrees of freedom of 19, we get a *p-value* smaller than 0.00001. Thus, with confidence of 95%, we reject the null hypothesis. In other words, there is a significant difference between the grades of the pre-test and the post-test exams of the group. We repeated the exam for the second group and obtained a *t-value* of 8.11. Using this *t-value* and degrees of freedom of 19, we get a *p-value* smaller than 0.00001. Thus, similar to the case of the first group, with confidence of 95%, we reject the null hypothesis. In other words, there is a significant difference between the grades of the pre-test and the post-test exams of the group.

7.2 Evaluation of the Effect of Artificial Word Generation

Wilcoxon *rank sum* test was applied to compare the grades of the post-test exams of the two groups. The null hypothesis H_1 was that there is no difference in the median between the two samples. The test was performed using Real Statistics [18] and repeated using MathCracker [20]. The resulting *Z-score* was -2.11 and the *p-value* was 0.0349. Thus, with confidence of 95%, we reject the null hypothesis. In other words, there is a significant difference in the median between the two samples.

7.3 Discussion of Results

From the results of the empirical evaluation, it is clear that ART had a positive effect on tutoring the students with and without artificial word generation. Nevertheless, students who were trained using artificial word generation scored higher grades in comparison to the other group. This proves the effectiveness of this novel technique and suggests it as a promising technique in language tutoring.

8 Conclusion

This paper presented ART, an ITS of Arabic word root extraction. ART is a cognitive tutor intended to teach students production rules needed for the extraction processes. ART works in two modes. In the passive mode, the student supplies the input word and ART responds by generating the root, the formal production rule and the simplified production rule used in the root generation. In the active mode, on the other hand, ART generates a word prompting the student to provide its root.

To help in tutoring the student, ART integrates several techniques in the active mode. Positive feedback is generated in case of a correct answer to motivate the student to continue learning and improving. On the other hand, in case of an error, a gentle negative feedback is generated. In this latter case, PAS is employed, where the student is guided to detect the error through scaffolding and self-explanation. If the error persists, the correct root is generated with the formal production rule and the simplified production rule used in the root generation. In addition to generating real words, ART generates artificial words to make sure the student can apply the production rules to unfamiliar words and that he/she does not merely memorize the roots of familiar words. Empirical evaluation has shown the effectiveness of the tutoring criterion of ART and suggests artificial word generation as a novel promising technique in language tutoring.

As a future work, ART will be extended to work in the integrated mode in which the student supplies the input word, but the system tutors him/her similar to the active mode. It will then accept input words with diacritics such as the *shadda*. We also intend to extend ART to cover additional production rules for Arabic word root extraction. When the number and types of production rules in ART increase, they will be classified into categories of related rules and the student will be able to select the required category for self-paced learning. Besides, ART is currently being developed

into an advanced ITS, which can model the student knowledge so as to automatically generate problems relevant to the student level of knowledge.

9 References

[1] Mitrovic, A., Koedinger, K. and Martin, B. (2003). A comparative analysis of cognitive tutoring and constraint-based modeling. 9th International Conference on User Modeling, Johnstown, PA, USA, pp. 313-322.

[2] Tepperman, J., Lee, S., Narayanan, S. and Alwan, A. (2011). A generative student model for scoring word reading skills. IEEE Transactions on Audio, Speech, and Language Processing, 19(2): 348-360. https://doi.org/10.1109/TASL.2010.2047812

[3] Sosnovsky, S., Dolog, P., Henze, N., Brusilovsky, P. and Nejdl, W. (2007). Translation of overlay models of student knowledge for relative domains based on domain ontology mapping. 13th International Conference on Artificial Intelligence in Education, pp. 289-296.

[4] Spada, H., Stumpf, M. and Opwis, K. (1989). The constructive process of knowledge acquisition: Student modeling. 2nd International Conference on Computer Assisted Learning, Dallas, TX, USA, pp. 486-499. https://doi.org/10.1007/3-540-51142-3_81

[5] Ohlsson, S. and Mitrovic, A. (2006). Constraint-based knowledge representation for individualized instruction. Computer Science and Information Systems, 3:1-22. https://doi.org/10.2298/CSIS0601001S

[6] Corbett, A., Kauffman, L., Maclaren, B., Wagner, A. and Jones, E. (2010). A cognitive tutor for genetics problem solving: Learning gains and student modeling. Journal of Educational Computing Research, 24(2):219-239. https://doi.org/10.2190/EC.42.2.e

[7] Andersson, J. and Johansson, H. (2015). Using clustering in a cognitive tutor to identify mathematical misconceptions. Master Thesis, Lund University.

[8] Mitrovic, A., Ohlsson, S. and Barrow, D. (2013). The effect of positive feedback in a constraint-based intelligent tutoring system. Computers & Education, 60:264-272. https://doi.org/10.1016/j.compedu.2012.07.002

[9] Ferreira, A., Moore, J. and Mellish, C. (2007). A study of feedback strategies in foreign language classrooms and tutorials with implications for intelligent computer-assisted language learning systems. International Journal of Artificial Intelligence in Education, 17:389-422.

[10] Dabbagh, N. (2003). Scaffolding: An important teacher competency in online learning. TechTrends, 47(2):39-44. https://doi.org/10.1007/BF02763424

[11] Mitrovic, A. (2005). The effect of explaining on learning: A case study with a data normalization tutor. 12th International Conference on Artificial Intelligence in Education, University of Amsterdam, The Netherlands.

[12] Al-Shalabi, R. and Evens, M. (1998). A computational morphology system for Arabic. Workshop on Computational Approaches to Semitic Languages, Montreal, Canada, pp. 66-72. https://doi.org/10.3115/1621753.1621765

[13] Abu Hawas, F. (2013). Exploit relations between the word letters and their placement in the word for Arabic root extraction. Computer Science, 14(2):327-341. https://doi.org/10.7494/csci.2012.14.2.327

[14] Yousef, N., Abu-Errub, A., Odeh, A. and Khafajeh, H. (2014). An improved Arabic word's roots extraction method using N-gram technique. Journal of Computer Science, 10(4):716-719. https://doi.org/10.3844/jcssp.2014.716.719

[15] Kanaan, R. and Kanaan, G. (2014). An improved algorithm for the extraction of triliteral Arabic roots. European Scientific Journal, 10(3):346-355.

[16] Al-Shawakfa, E., Al-Badarneh, A., Shatnawi, S., Al-Rabab'ah, K. and Bani-Ismail, B. A comparison study of some Arabic root finding algorithms. Journal of the American Society for Information Science and Technology, 61(5):1015-1024. https://doi.org/10.1002/asi.21301

[17] Elazhary, H., Alharthi, A., Balkhi, E., Aljahdali, G., Zagzoog, D. and Alkhammsh, A. (2015). Automated tutoring of Arabic word root extraction. International Journal of Scientific & Engineering Research, 6(7):687-691. https://doi.org/10.14299/ijser.2015.07.002

[18] Real Statistics Using Excel, http://www.real-statistics.com, [Online; accessed: 2016-07-01].

[19] Social Science Statistics, http://www.socscistatistics.com, [Online; accessed: 2016-07-01].

[20] Free Math Help, http://www.mathcracker.com, [Online; accessed: 2016-07-01].

10   Authors

Hanan Elazhary earned her B.Sc. and M.Sc. degrees from the Department of Electronics and Communications Engineering, Cairo University. She earned her Ph.D. degree in Computer Science and Engineering from the University of Connecticut, USA. Currently, she is an associate professor in the Computer Science Department, Faculty of Computing & Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia and the Computers and Systems Department, Electronics Research Institute, Cairo, Egypt. Here research interests include distributed systems, software engineering and intelligent tutoring systems.

Nabila Khodeir is a researcher in the Informatics department at the Electronics Research Institute, Cairo, Egypt. Her research interests include intelligent tutoring systems, user modelling and natural language processing. She earned her Ph.D. and ME from the Electronics and Communications department at Cairo University.