Journal of Intelligent & Robotic Systems (2021) 102:48
https://doi.org/10.1007/s10846-021-01413-2

REGULAR PAPER

Design and Implementation of a Robotic Architecture for Adaptive Teaching: a Case Study on Iranian Sign Language

Salar Basiri [1] · Alireza Taheri [1] · Ali Meghdari [1,2] · Minoo Alemi [1,3]

Received: 25 August 2020 / Accepted: 4 May 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
Social robots may soon be able to play an important role in expanding communication with the deaf. Based on the literature, adaptive user interfaces lead to greater user acceptance and increased teaching efficiency compared to non-adaptive ones. In this paper, we build a robotic architecture able to simultaneously adjust a robot's teaching parameters according to both the user's past and present performance, adapt the content of the training, and then implement it on the RASA robot to teach sign language based on these parameters in a manner similar to a human teacher. To do this, a word to teach in sign language, repetition, speed, and emotional valence were chosen to be adaptive using a fuzzy logic mechanism. Then, two groups of participants were recruited. For the first group, the robot teaches without the adaptive architecture, while for the second group, the teaching is done with the adaptive architecture. The assessment phase was conducted with 8 users in person and 48 users virtually. A standard UTAUT questionnaire was selected to assess the effectiveness of this methodology by comparing different items from the two groups of users. Statistical analysis of the T-test and Cohen's d effect size found that the second group felt the robot's adaptability significantly more than the first group, indicating that the methodology used in this study was effective and that the robot's ability to adapt was felt by users. In addition, the results of the two groups were significantly different in several other items, revealing the effects of the adaptive architecture.

Keywords Social robots · Adaptive teaching · Fuzzy logic · Human-robot interaction (HRI) · Iranian sign language (ISL) · Intelligent tutoring systems (ITS)

1 Introduction

As robots move from industrial environments to homes, studying and optimizing Human-Robot Interactions (HRI) has become very important. While users in industrial environments are compatible and comfortable with the features of robotic equipment used in their activities, the same degree of familiarity is missing in interactions with social robots. This issue of technical acceptability is especially important for the future success of social robots and is studied in the field of social interactions [1]. In recent years, research results have been satisfactory for using human-social robots in various fields such as education, medical treatment, entertainment, etc., i.e., using social robots to teach Sign Language (SL) to children with hearing impairments [2]. Research has shown that using a social robot in the process of educating children can greatly increase their educational efficiency, and according to child psychological research, when used as a game, can be an asset in their intellectual development [3]. The issue of adapting a robot's capability to users in the field of human-robot interaction has recently received attention in robotic studies.

It has been well demonstrated in the literature that adaptive user relationships lead to significantly greater robot acceptance by users than non-compatible robots [4]. Adapting social robot teaching methods to the user and/or training content improves the robot's perceptual and cognitive level. In our case, this ability to adapt can push the process of teaching sign language to new levels of efficiency and convenience, thereby helping to build better social relationships with hearing people for hearing impaired and deaf individuals.

✉ Alireza Taheri
artaheri@sharif.edu

1 Social and Cognitive Robotics Laboratory, Sharif University of Technology, Tehran, Iran

2 Chancellor of Fereshtegaan International Branch, Islamic Azad University, Tehran, Iran

3 Faculty of Humanities, West Tehran Branch, Islamic Azad University, Tehran, Iran

Published online: 27 May 2021

Inspired by classic papers that deal with adaptable interfaces ([4, 5]), as well as works done in [1, 6], previous works can be classified into three categories. The categories were based on the robot's mechanism of adaptability and user profiles, whether or not this process is done by building user models, and if so, whether the user model is static or dynamic (i.e., stability and continuity of user models over time). Systems with a dynamic model continuously improve their perspective and overall view of the user over time and adapt to their changes, which calls for increasing the autonomy and robustness of the system. The adaptive teaching process requires a model of the user's personality, preferences, and performance and needs to be updated as these factors change over time. Therefore, we only review the state of the art robotic systems that use a dynamic user model to adapt their system to users.

Karami et al. addressed the issue of user adaptation as the primary objective of a mobile robotic system [7–9]. The system learns user preferences using interactive effects delivered during repeated human-robot interactions, and then uses a Markov Decision Process (MDP) decision-making formulation to review its policy according to the user preferences. One of the results highlighted in the work is that the robot can adapt very quickly to a new user by using the model to generalize the gained rewards, as it has experienced interactions with other users in the past. A "pro-active" system provided by Grosinger et al. [10] integrates into the ACCOMPANY project introduced by Abdollahian et al. [11]. The system maintains a user mode, as well as a set of rules for evolving that mode overtime to keep the user in a "good" mode. By observing and analyzing the user's choices as well as the environment, the robot identifies opportunities for appropriate action that can distract the user from reaching an unfavorable situation. Ali and Tapus conducted a study [12] that used a new fuzzy methodology to identify user emotions online. Both Nao and Alice robots were used for functional testing. An attempt was made to adapt the robot's behavior to build a better long-term interaction with the user based on his personality and emotions. Westland, Gordon, et al. [13] introduced a teacher assistant companion robot to help English-speaking children to learn a second language (Spanish). Two aspects of this interaction were personalized for each child: 1) The content of the game (i.e., choosing which words to teach), and 2) The robot's emotional response to the child's emotional state and performance. Then, in the same year, Gordon and Breazeal [14] and the authors of [15] in 2016, introduced an Intelligent Tutoring System (ITS) that aims to help children learn to read. The system maintains knowledge of the user's reading level, which is evaluated and updated periodically with an active learning technique. This information is then used to adapt the game that the child and the robot are playing, and the robot also adjusts its motivational strategy using verbal and non-verbal activities. Aylett et. Al. in [16] also introduced an empathic robot to help the user learn geography, it keeps a record of the person's skill level, such as how to use polarization and map symbols, and adapts its actions to this level of skill. This practice occurs in the organized literature of an architecture called EMOTE, and the process of robotic perception and adaptability occurs in a modular manner. Tozadore in [17] also implemented an adaptive cognitive architecture called R-Kessel to provide a new mechanism to help improve interactive teaching activities (such as teaching geometric shapes to children) on the Nao social robot. In addition, in [18] an adaptive model of educational resources is presented that promotes robotic teaching in different courses. Based on the recorded student data, the system builds student models that evolve over time and categorize users based on their skill levels and knowledge. In addition, the learning environment is dynamically adapted to each user according to the actions taken.

Since a great deal of the mechanisms proposed in the literature function directly based on the data derived from the user/environment, they usually have to cope with the issue of outliers in the data and/or uncertainties in the modeling process. Therefore, the output stability of these systems is proportional to the degree that these outliers and uncertainties are handled. For example, the authors in [19] investigated the stabilization problem of neural networks with unbounded continuously distributed delays via impulsive control. By establishing an impulsive infinite delay differential inequality, they derived some sufficient conditions ensuring the stabilization of the unique equilibrium point. Similarly, in [20] the authors apply techniques to calculate filtering parameters that guarantee the finite-time boundness and strict dissipativity of the filtering error dynamic system. Also, in [21], the authors addressed the issue of fault-tolerant control in systems with different sensors by proposing two strategies for robust estimation of linear stochastic models in presence of model oversimplifications and noise. Also in the field of controls, the work [22] examines a PD-type iterative learning control algorithm for a class of discrete spatially interconnected systems with unstructured uncertainty. Although the methodology of this work shows promising results in practice, it is adversely affected confronting systems with unmeasurable states. Uncertainties in fuzzy control systems have also been discussed in the literature. For example, in [23] the problem of various data missing in the design of a fuzzy controller in networked control systems is addressed. The authors used an auxiliary random series method to describe the data transferring in the network. By theory analysis and simulation, the compensation of missing data implemented by a buffer showed to be effective.

In this study, we designed and implemented an adaptive teaching architecture to empower the RASA social robot for adaptive teaching of Iranian Sign Language (ISL) to users. This fuzzy logic based architecture makes it possible to adapt the robot's teaching output using four aspects based on the users' past and present performance: 1) What words should be selected for training, 2) How many times each word should be repeated in the training process, 3) What should be the speed of the performance of SL signs, and 4) What should be the emotional reaction (valence) of the robot to the user's performance? A very important innovation and the challenge of this research is that the robot's compatibility and adaptation to the users is maintained through performance feedback and is simultaneously accomplished through two features: one, the robot's logic concerning its general teaching program (General adaptability), and two, the personalization of the teaching program by tailoring it to a particular user (specific adaptability). After each training session, the parameters governing the adaptability system get modified both in the user profiles and word profiles at the same time, which enables robot teaching program to be adaptive toward a specific user AND all of the following users. By executing this scenario: a) During teaching sessions, the robot's inner teaching logic will be adapted, which will make it faster to adapt to new users, and the robot will gain experience from its previous training, and b) When interacting with the same user again, the robot teaches more intelligently, based on the user's previous and current performance.

As far as the authors know, there has been no such research on making the robot's teaching program both generally and specifically adaptive based on the users' performance at the same time. As seen in previous works, most systems start building a dynamic user model for new users from scratch. They initially create a match between the new user and previous users and then select the closest model to them. This slows down the process of adapting to the new user. In this work, all previous teaching will enhance the robot's ability to adapt, and new users will benefit from the robot's past teaching experience.

In section 2, the methodology of this work is presented, the robotic architecture implemented is described, and each block is examined in details. Then, the experimental setup and the evaluation scenario is discussed in detail. In section 3, the results of the experiment and the statistical parameters are presented, and a discussion is made based on the results. In section 4, the limitations of this study and suggested future work are presented. Modeling errors and uncertainties of the algorithm are discussed in section 5. Finally, in section 6, a conclusion is made on the proposed methodology.

2 Methodology

To make the process of robotic teaching adaptive, it is necessary to form two different groups of dynamic classifications, 1) Word profiles and 2) User profiles, and have them evolve over time. In the classes related to the word profiles, the following cases are used to adjust the degree of truth (membership) of the word in a set of fuzzy membership functions (i.e., easy, medium, and hard word): 1) the average number of times performed by people to achieve the correct performance, 2) the average accuracy of people's performance, and 3) the average speed of each person's performance. Similarly, the same parameters are utilized in the user profile, only this time they are related to that particular person, to adjust the user's degree of truth in a set of fuzzy membership functions (i.e., weak, medium, and strong user). Therefore, the four adaptive output parameters of the robot will be a mathematical function of both the class parameters of the word itself and the parameters of the user profile at the same time. Linking the mathematical parameters of these two classes based on the individual's performance parameters to the robot outputs was one of the most important challenges.

2.1 Fuzzy Logic

The term fuzzy logic was first coined in 1965 by Professor Lotfizadeh in the theory of fuzzy sets [24]. Since the introduction of fuzzy logic, this tool has been used in many applications of engineering and artificial intelligence. In the field of robotics, fuzzy logic has been used in cognitive processes and artificial intelligence [25–27], and in the field of control, it has been used in many applications such as adaptive fuzzy control [28, 29]. Fuzzy logic was formed based on the observation that humans naturally make their decisions using information that is inaccurate and even ambiguous. This logic is a mathematical tool by which this ambiguity and inaccuracy in data can be expressed. Fuzzy models can identify, express, manipulate, interpret, and make use of data in which there is a degree of uncertainty and ambiguity [30]. While variables are described with numerical values in classical mathematics, non-numerical and so-called linguistic values are used to describe rules and variables in fuzzy mathematics [31].

Teaching is a comparative process that deals with many linguistic variables [32, 33]. When reviewing the input information received from students or the content of a lesson, a human teacher does not process the data with classic crisp logic; but instead works with linguistic variables; for example, he/she generally knows in his/her mind that one user is weak, another user is average, and another user is strong. The same is true for the teaching content; some lessons are difficult, some are medium, and some are easy. Therefore, a change in the average score of a person does not suddenly change his judgment about that person, but rather, it changes with a gentle and reasonable slope. Moreover, this logic helps the teacher adjust the teaching program easier and faster in terms of language variables. For example, when a student is weak and the content is difficult, the teacher needs to slow down and increase repetition; conversely, when the lesson is easy and the user is strong, the teacher can teach quicker and repeat less. In order for a teaching robot to perform similarly to a human teacher, it is necessary to work with linguistic variables instead of crisp variables. This makes the adaptability process more efficient, the output of the robot more compatible with inputs, and the logic with which the robot works more descriptive for humans. In this paper, the robot's fuzzy databases were created for users and words, and then the robot's adaptive logic was implemented by building the robot's fuzzy rules.

2.2 RASA Robot and the Glove Sensors

The RASA robot is a humanoid robot with a total of 32 degrees of freedom (29 degrees of body movement including active fingers and 3 degrees of lower-body movement); it was designed to teach Iranian Sign Language to deaf children [34]. The robot receives input data from users' performance using a glove, processes it with a deep neural network, and then communicates with children to teach them different ISL signs. Figure 1 shows the robot during ISL gestures. The Glove sensor used in this study was a Perception Neuron V2 suit from Noitom Ltd. It utilizes some IMU and gyro sensors to capture the finger and hand movements.

2.3 Adaptive Teaching Architecture

In general, to get a robot to accurately perform a particular task in different environmental conditions, it is good practice to construct a general architecture that considers which components do each small part of the work and work well together. To do this, the design of the required algorithm is first simplified, and then a step-by-step design can be used to determine (accurately and effectively) which part is not performing its role well. Figure 2 shows the general architecture designed for the adaptive ISL teaching presented in this work.

We tried to design the architecture so that it includes all the functions a robot requires to be able to teach sign language, similar to a human teacher. In other words, if all the components of the proposed architecture work and communicate with each other properly, the process of teaching is done adaptively and is consistently in line with the user's performance and the necessary teaching content, and the main goal of this research is satisfied. We now turn to the algorithm involved in this architecture.

Initially, the username is given to the algorithm as an input. The training process begins by simply giving the name/ID of the user to the robot. No other input is required. Then, in the user Scanner Module, the robot checks the user's history to determine the student's weak and strong words. The Word Selector Module then determines what words should be selected for the current training based on the information successfully passed in the previous module. The Word Analyzer Module reads the information recorded in the word profile and uses the information to calculate the word score according to the word history. The User Analyzer Module does the same for the user's profile.

Next, we come to one of the main and largest modules of this architecture, namely the Robot Output Generator Module. This module is the core of adapting the robot's teaching process. By taking the user's profile and the word profile information, it calculates the user's and the word's scores, adjusts their position in the fuzzy membership functions, and determines how the robot's outputs should change. The Advertiser then codes and sends the emotional response, number of repetitions, and speed of performance to the robot. Now the robot performs the Word Selected for session according to the calculated output parameters and waits for the user. This waiting time is related to the user's readiness and the duration of the robot's performance to teach the appropriate words and is determined by the start test module. A supervisor helps the start test module to initiate. Now, it is up to the user to use the glove to perform the presented word, which the robot then evaluates as successful or not.

The Data Recording Module stores the data transmitted from the data collection glove. The Sign Recognition Module determines what word the user has performed and its degree of accuracy using the position data of the user's fingers/arms captured from the glove. It then sends its output back to the robot to set the robot's emotional state; this cycle is repeated in an internal loop as many times as necessary determined by previous modules. When the internal loop is

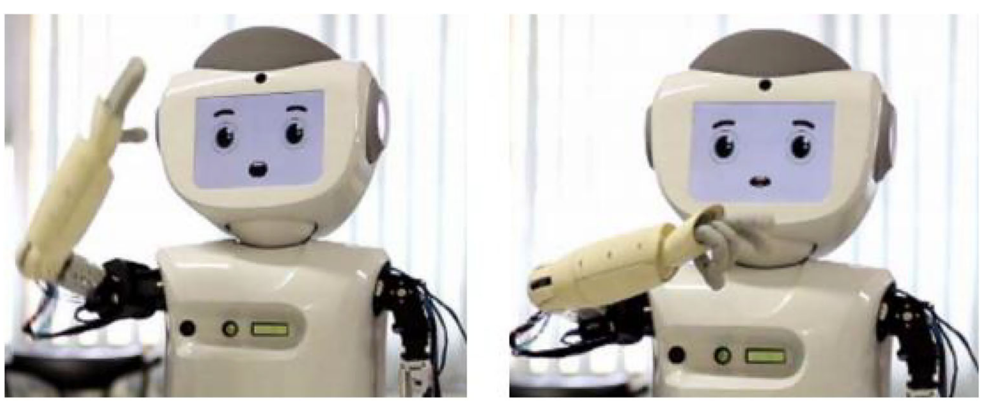

Fig. 1 The RASA robot performing ISL gestures

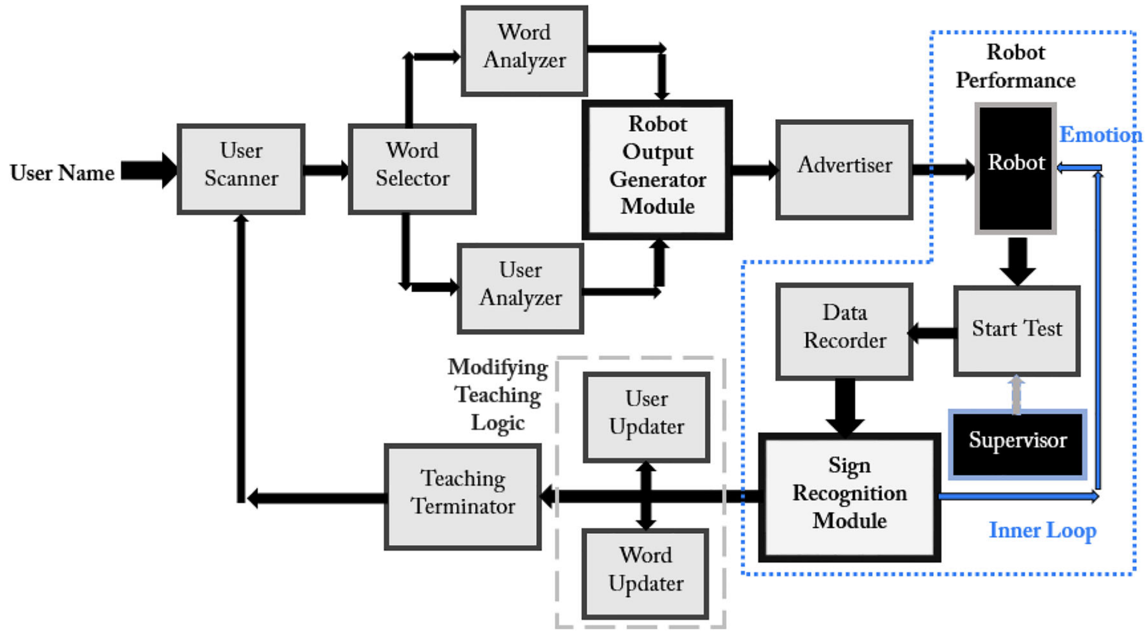

Fig. 2 The adaptive teaching architecture proposed in this study

complete, the Updater Modules write the new information on the user's and the word's profile. The Teaching Terminator Module decides whether the training process should continue for that user and if the main (outer) teaching loop needs to be repeated or the training process can be terminated for that user. Correctly implementing this architecture shows how a robot can adapt its teaching process in a manner somewhat similar to a human teacher.

We examine some of the important components of this architecture in more detail in the following sections.

2.3.1 Word Selector Module

This module performs one of the important tasks of the algorithm. We claimed that the robot adapts to its user by customizing the words it teaches. The words chosen for training depend on what words the user has successfully passed or failed in their previous interactions with the robot. Therefore, this module uses a random selection process with a heterogeneous probability density to select the words. In this way, when that word is rejected or passed, the weight of the probability of choosing it changes, but the weight of all the words before training for the user is equal to 1. In other words, if we show the weight of the probability of choosing the n_{th} word with $weight_n$ and the number of passes and rejections for that word with n_p and n_f, respectively, we obtain the following equation:

$$weight_n \leftarrow (1-0.5n_p+0.5n_f)\mathbb{I}_{(0<weight_n\leq 3)}+0.1\mathbb{I}_{(weight_n\leq 0)}+0.3\mathbb{I}_{(3<weight_n)}$$

Where:

$$\mathbb{I}_{(Statement)}=\{1|Statement\}$$

Is the identity function that returns 1 only if the statement is true, and 0 otherwise.

In other words, words that the user has previously been able to pass are not completely removed from the tutorial cycle but are less likely to be selected in a random process. This makes it possible for the user to retain those words for future tutorials, allowing them to be reviewed and consequently keeping them in mind. Therefore, in practice, the teaching process is unique to each user and focuses more on the words in which the user is weaker.

2.3.2 Analyzer Modules

At this point in the algorithm, we know which user and which word to teach. Thus, the Word Analyzer and the User Analyzer modules open the word and user profiles, respectively, to extract information about the user's past performance and perform calculations accordingly. These modules read information from all training sessions, averaging three parameters: 1) number of repetitions, 2) accuracy, and 3) speed of the user in all previous sessions. That is, a total of six parameters are sent as output to the next module: three of which are related to the "user" and three of which are related to the "word". These parameters will be used in the next module to calculate the score. If the user or word is not used in any training session, their score is 50 (out of 100), and their status is considered perfectly medium/normal.

2.3.3 Robot Output Generator Module

This module forms one of the most vital points of the teaching algorithm as the principle of adapting the robot's teaching occurs here. In this block, using the six inputs received from the previous modules, the output of the robot is determined so that it is adaptive for that training session (i.e., that particular user and that particular word). The six inputs of the module are: 1) the average accuracy of the user in performing the signs ($\overline{Acu_u}$), 2) the average number of the user repetition in performing words $\overline{Rep_u}$), 3) the average speed of the user's performance in different words ($\overline{Spd_u}$), 4) the average accuracy of different users on that word ($\overline{Acu_w}$), 5) the average number of times users need to repeat before passing that word ($\overline{Rep_w}$), and finally, 6) the average speed of users performing that word ($\overline{Spd_w}$). The three outputs of this module are 1) the number of times the robot has to perform the word for the user (Rep_R), 2) the robot's emotional reaction to the user's performance (Emo_R), and 3) the speed at which the robot must perform the word (Spd_R). Figure 3 shows a schematic of the internal blocks of this module.

As seen, three parameters related to the word profile are sent to the internal module to calculate the word score, and three parameters related to the user profile are sent to the internal module to calculate the user score. Depending on the parameters given, these modules assign a score between 0 and 100 to the user and the word. Then, using the given score and the predefined fuzzy membership functions, the degree of truth (membership) for the user and the word to their fuzzy classes is determined. The fuzzy database is defined in these two modules. Then, with the degree of truth of the fuzzy functions, the Output Regulator Module uses the fuzzy rule base to determine the defuzzified outputs of the robot. Therefore, this module calculates the robot's adaptive outputs by taking the user's performance parameters. To make the point clearer, let us take a closer look at some of these internal blocks.

Word Score Calculation Module By capturing and averaging the parameters of repetition number, accuracy, and the performance speed of users on the word, the goal of this module is to determine the word's degree of truth for the "easy" class (α_w), the "medium" class (β_w), and the "hard" class (γ_w). To do this, it must first assign a score to the word using the parameters, and then, by using the position of the score in the fuzzy membership functions, determine the three parameters related to the three fuzzy classes: easy, medium, and hard.

To score the word, this module uses a non-linear interpolation function. That is, a human teacher first determines the score of the word against a certain number of input parameters and then uses the interpolation for other inputs. The word score (R_w) is a score between 0 and 100, with 0 for quite easy and 100 for quite hard. The word score is determined by a second-order interpolation between points assigned by the human teacher. Then, based on the fuzzy database of words, the degree of truth for the fuzzy classes of the word is determined. The shape of these fuzzy membership functions is shown in Fig. 4.

Therefore, by defining fuzzy membership functions linearly according to the figure above, the degree to which the word belongs to each of the fuzzy classes is easily determined via the following equation:

$$\alpha_w = \mathbb{I}_{(R\leq 25)} + \frac{50-R}{25}\mathbb{I}_{(25<R<50)}$$

$$\beta_w = \frac{R-25}{25}\mathbb{I}_{(25\leq R<50)} + \frac{75-R}{25}\mathbb{I}_{(50\leq R\leq 75)}$$

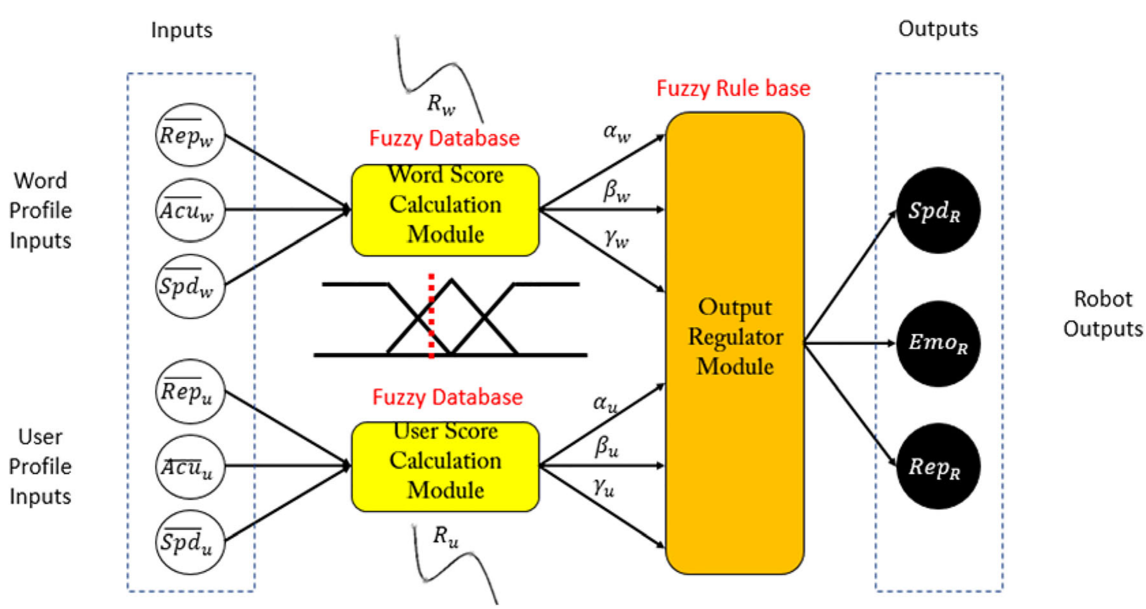

Fig. 3 The internal blocks of the Robot Output Generator Module

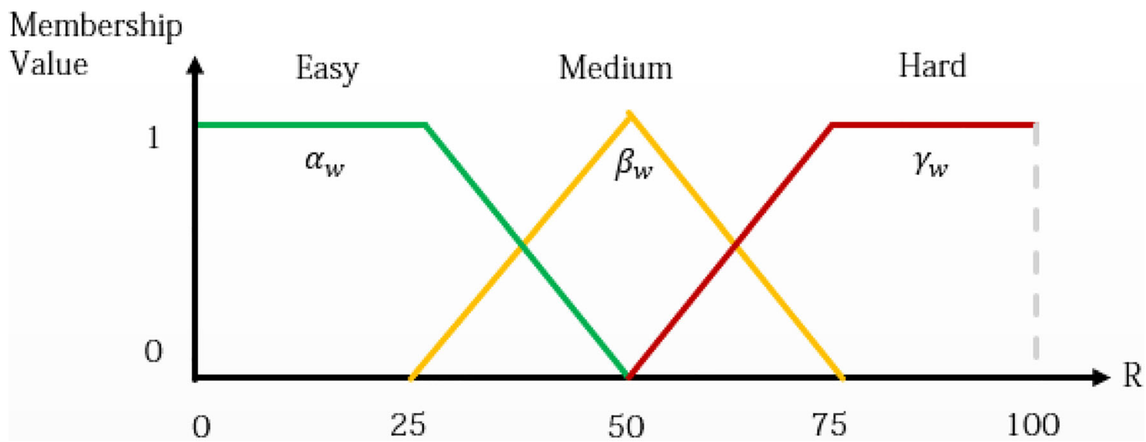

Fig. 4 The fuzzy membership functions related to the words

$$\gamma_w = \frac{R-50}{25}\mathbb{I}_{(50<R<75)} + \mathbb{I}_{(75\leq R)}$$

This is where the fuzzy databases are built. According to the above equation, scores below 25 mean quite easy and above 75 mean quite hard. Scores between these two values are considered to belong to two fuzzy membership functions, i.e., a score of 60 means that the word is both medium and difficult, but we believe it is more medium. The linearity of the membership functions speeds up the robot's performance significantly and allows real-time adaptations.

The User Score Calculation Module does the same thing with the user profile. That is, a human teacher assigns a score to a user based on some inputs and performs interpolation for the rest of the inputs to calculate the user's score. The user's fuzzy classes include "weak" (α_u), "medium" (β_u), and "strong" (γ_u). Similar to the words, user membership functions are considered linear. We need this degree of fuzzy memberships at the fuzzy rule base to calculate the robot's output.

Output Regulator Module Using fuzzy membership degrees for each class, this module calculates the degree of truth for each of the rules in the fuzzy rule base, and then uses their weighted average to generate the appropriate output of the robot. The fuzzy rule base is formed for the robot's output speed, the number of robot repetitions, and its emotional response (in both failed and passed situations). The degree of truth for any fuzzy rule is equal to the product of the multiplicity of membership degrees in each of its components.

To make it easier to show fuzzy rules, we have compiled them into nine celled tables that represents the nine fuzzy rules. Figure 5 shows the fuzzy rules related to the number of repetitions and the output speed of the robot. Figure 6 relates to the fuzzy rules of the robot's emotional response in the correct performance mode and the wrong performance mode.

Figure 5 states that if the user status is normal and the word status is difficult, the number of repetitions of the robot should be high and its performance speed low. Conversely, when the word is normal but the user is strong, the number of repetitions should be low and the performance speed should be high. The belief is that each of the nine fuzzy rules will be the multiplication of the corresponding truth degrees of the word and the user class. Finally, when the belief in all nine rules is taken into account, their average is taken to perform de-fuzzification operations, and a numerical value is calculated for the repetition number and speed of the robot. The same goes for the robot's emotional response. For example, if a weak user is unable to pass a hard word, the robot's emotional response is happiness to increase the user's motivation. On the other hand, a strong user can successfully pass a medium word, so the robot's emotional response is neutral.

Following this description of the Robot Output Generator Module, we can see how the output of the robot (hopefully) adapts to the user's performance, as well as the difficulty or simplicity of the teaching content. After all the robot's output parameters have been calculated, these parameters must be transferred to the next module so that the robot can finally perform using these parameters.

2.3.4 Sign Recognition Module

The function of the Sign Recognition Module is to determine how accurately the user has previously performed, based on the data (the user's hand positions) saved from the previous module. By recording the position of the hand over time, this module should be able to deliver an output list determined by what belief the user has performed for each of the available words. This module was developed in our previous work (currently under review) with the help of a Deep Neural Network (DNN) and the state-image methodology. We observed accuracy of 99.7% for the proposed DNN for a set of 15 ISL signs.

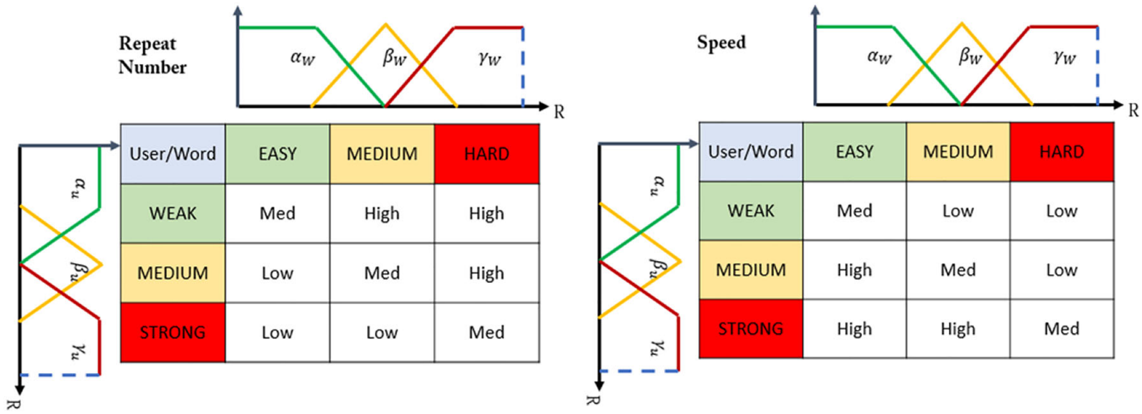

Fig. 5 The fuzzy rules related to the speed and the number of repetition outputs

2.3.5 The Inner Loop of Teaching

If we take another look at the main (outer) teaching loop in Fig. 2, it is clear that there is also an inner training loop marked with a blue box. The inner teaching loop is repeated as many times as calculated in the Robot Output Generator Module. Therefore, every time the robot performs a word, it waits for the user to perform as well, then evaluates the user's response, and repeats the training by reacting emotionally with the user. This process is repeated as many times as relative to the user's level and the word's level. This loop is then completed and the next modules perform their operations.

2.3.6 Initial Conditions

Now that the adaptive teaching system has been clearly explained, it is good practice to address what initial conditions affect the system's behavior. There are several initial conditions that need to be determined before the system can run. The first and maybe the most important initial conditions are in the Word/User score calculation module, where these modules have to assign a score to the user/word based on their average performance parameters. This is done by performing a polynomial data-fitting and interpolation on a set of limited data points that are set once by a human user. For example, a human teacher has determined that an average repeat number of α, an average accuracy of β, and average speed of γ gives a score of s. Changing this initial data will change the whole performance of the adaptive system significantly.

There is also an initial condition in the word selector module that determines how often the robot urges a repeat of the words failed by the user. This initial condition is actually the set of probability weights assigned to the failed/passed words.

Lastly, there are initial conditions in the Robot Output Generator Module that determine the shape and the location of fuzzy membership functions and the format of the fuzzy rule base. While linear membership functions are used, other functions such as Gaussian functions can be implemented, and the parameters of these functions can also be modified as initial conditions. In the fuzzy rule base, the fuzzy output of different sets of fuzzy inputs can also change. For example, the fuzzy emotional reaction of the robot can be set differently as shown in Fig. 6.

2.3.7 Computational Burden

The computational burden of the algorithm takes place mainly at the Robot Output Generator Module and the Sign

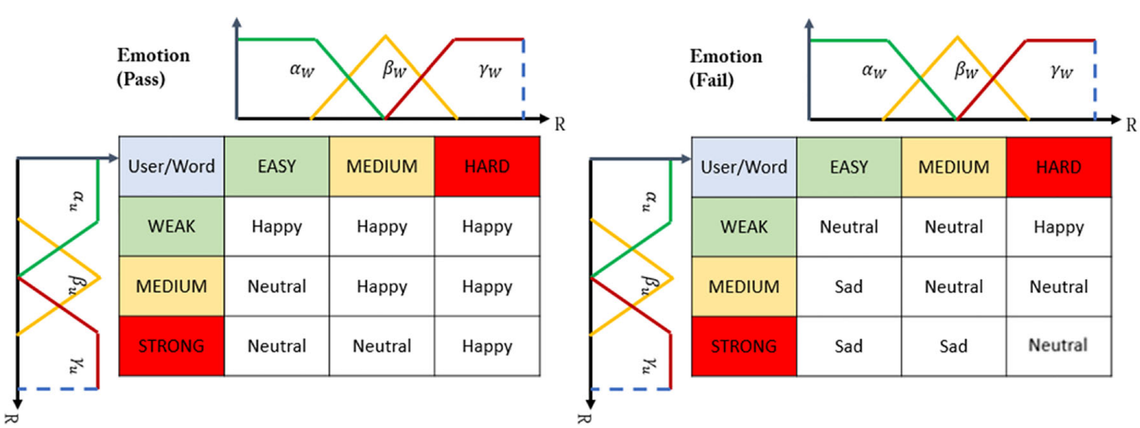

Fig. 6 The fuzzy rules related to the robot's emotional reaction when user fails/passes

Recognition Module. Only simple R/W and mathematical operations are conducted in the other modules. In the output generator, only at the first run, an optimization process occurs to initialize the fitting equation of the Score Calculation Modules. Then, simple linear arithmetic operations are needed in the fuzzy database and rule base to calculate the output. In the Sign Recognition Module, the deep neural net is pretrained. Therefore, the only computational cost happens when the new data is fed to the deep neural network to generate the output. Since no training is needed when new data arrives, there is no heavy computational burden at this stage, making it appropriate for real-time applications.

2.4 Implementing on RASA

To complete the last block of the adaptive teaching architecture, we expect that the robot will respond appropriately after receiving input signals from the system. The architecture sends different variables to the robot, each of which triggers a specific function in the robot. These variables are:

1) The word that the robot should perform: referring to its list of learned vocabulary, the robot finds the spatial coordinates of its arms, and commands the actuators to move the arms in a certain path.
2) Robot's performance speed: this rate determines how fast the motors should work while at the same time maintaining a smooth and acceptable trajectory.
3) The robot's facial reaction: utilizing this parameter, the robot makes a sad, neutral, or happy face.
4) The robot's vocal reaction: it must announce what word is being taught, encourage the user if he/she performs correctly, and express sadness if they perform incorrectly.

2.5 Evaluation Scenario and the Participants

In this section, it is necessary to determine what the experiment conditions are and how the teaching scenario is implemented. First, an explanation of the criterion for passing or failing at the words are given. Then, we discuss the experimental setup, how to evaluate the process using questionnaires and statistical analyzes, and lastly, the effect of adaptation on teaching quality.

The adaptive architecture determines how many times a word must be performed for a user. If the number of times the user repeats the word to get a correct answer is less than or equal to the number of repetitions assigned to the robot, that word will be passed; otherwise, it will be rejected.

Two experiments were performed to evaluate the research: In one, the RASA robot taught a total of 8 human users in person and stored this training data. In the other, the RASA taught a total of 48 participants virtually and stored their data. In both experiments, the users were divided into two groups, and each person was taught for a fixed 10 min. In each experiment, the robot behavior was nonadaptive for the first group, while for the second group, the architecture was implemented and the robot was adaptive. A total of twenty sign language words exists in the vocabulary domain of RASA. Group members were randomly selected to minimize the effects of individual intelligence and memory on the group average. Roughly half of the total participants had robotic knowledge, while the other half did not. There was an approximately equal gender balance, and the age range was from twenty to fifty. As the robot became more experienced after each training session and corrected its view of the words, it was expected that the training of the adaptive group would be more effective than the nonadaptive one. In other words, since in each teaching session, the robot's inner adaptive parameters are regulated, it learns how to teach the following users more effectively. Hence, when faced with the second group, the robot was smarter and better able to adapt to them as well as the teaching content; therefore, the robot's ability to adapt can be assessed by comparing the two groups. Different items of a standard questionnaire filled in by the two groups of participants were compared to assess the effects of the adaptation. Several parameters were used to compare the effectiveness of the training process in the two groups.

For the first parameter, we considered the average number of the group's failed words in the first experiment. As a hypothesis in our HRI test, we expected that by implementing the proposed adaptive teaching, the participants who were trained in the adaptive group should be able to pass significantly more words because the robot knows how many times and how fast each word should be performed for them. In addition to assessing the effectiveness of the training in terms of the failed words, another assessment was also done by examining various items on standard questionnaires. After the training, the participants filled out the questionnaire in which, in addition to the effectiveness of the training, the quality of their interaction with the robot was also measured. By comparing the mean of the different items of the questionnaire for the two different groups in both experiments, we can examine the effect of adaptive training in different cognitive aspects and achieve useful results in the context of the human-robot interaction.

Various studies on socio-cognitive robots have provided questionnaires to assess a robot's acceptability [35]. Models have also been developed to evaluate various aspects of human-robot interaction; one of the most well-known is the Unified Theory of Acceptance and Use of Technology (UTAUT) [36]. After reviewing and studying the research and questionnaire done in [37], we decided to use a Persian translation of the UTAUT questionnaire to evaluate our work's acceptability. In addition to the standard items, the adaptability of the robot, and its effect on the user were also incorporated into the questions. Each participant used a 5-point Likert scale to respond to the questions. The questions for each item were in accordance with the UTAUT, with a few changes so that they can be used in this research. This questionnaire was based on the UTAUT since it integrates qualities of several leading technology acceptances, such as: Theory of Reasoned Action (TRA), Technology Acceptance Model (TAM), Theory of Planned Behavior (TPB), Combined TAM and TPB (C-TAM-TPB), Innovation Diffusion Theory (IDT), and Social Cognitive Theory (SCT). The items investigated in the UTAUT standard questionnaire are shown in Table 15. The questions used in this research are presented in Table 16.

Users gave a score of 1 to 5 to each of the above questions (i.e., 1: I totally disagree, 2: I disagree, 3: I neither agree nor disagree, 4: I agree, and 5: I totally agree). In each item, in addition to the overall conclusions on the statistical population, the training groups were compared, and then a statistical analysis of the T-test was performed on the results. Because the number of participants in the first study is not sufficient to draw general conclusions, in addition to the T-test, the Cohen's d effect size was also calculated for each item. In this way, and with this training scenario, the robot's ability to adapt was tested and preliminarily evaluated.

Two experiments were conducted to evaluate the effects of the proposed methodology. In the first experiment, a total of eight participants were taught by the robot in-person. It should be noted that due to the relatively small statistical population, the results of this experiment are used only to evaluate the performance of the implemented algorithm in the previous sections. These results are only for a preliminary exploratory assessment of the proposed algorithm, and to create a field of questions for more detailed clinical studies in larger independent statistical studies with the in-person target population. However, since the teaching in this experiment is performed in-person, the results can be useful to examine the real effects of the teaching, as the participants can interact with the robot in a real environment. Figure 7 shows a snapshot of the performed HRI setup, including a user wearing a sensory glove placed in front of the robot. The supervisor initiates and supervises the teaching process.

In the second experiment, a total of 48 participants were taught by the robot virtually. Due to a sufficiently large number of participants in the second experiment, the results of this experiment are of higher statistical validity. The participants were shown a video of the robot doing the teaching process, adaptive for half of the population and nonadaptive for the rest in separate videos. The videos are recorded in a first-person view, similar to sitting in front of the robot and interacting with it, to give the maximum sense of the robot teaching process to the users. Figure 8 shows the second experiment setup used in the assessment phase of the project. In the video, the robot performs several ISL signs, and the person viewing RASA performs accordingly, sometimes wrong and sometimes correctly, and watches the robot's reaction. We have tried our best to keep the similarities to the in-person experience as high as possible.

While the virtual experiment and assessment tried to be as close as possible to an in-person experiment, there is always a difference between having a real-world and virtual experience with a robot. Therefore, the results of both experiments are reported as complementary results of the assessment phase of this work, each having its particular advantages/disadvantages.

3 Results and Discussions

3.1 Experiment 1: Quality of the Teaching Based on the Training Data

We first compared the two groups of the users in terms of the number of failed words in the teaching process. This measure demonstrates how cleverly the robot is able to assess and

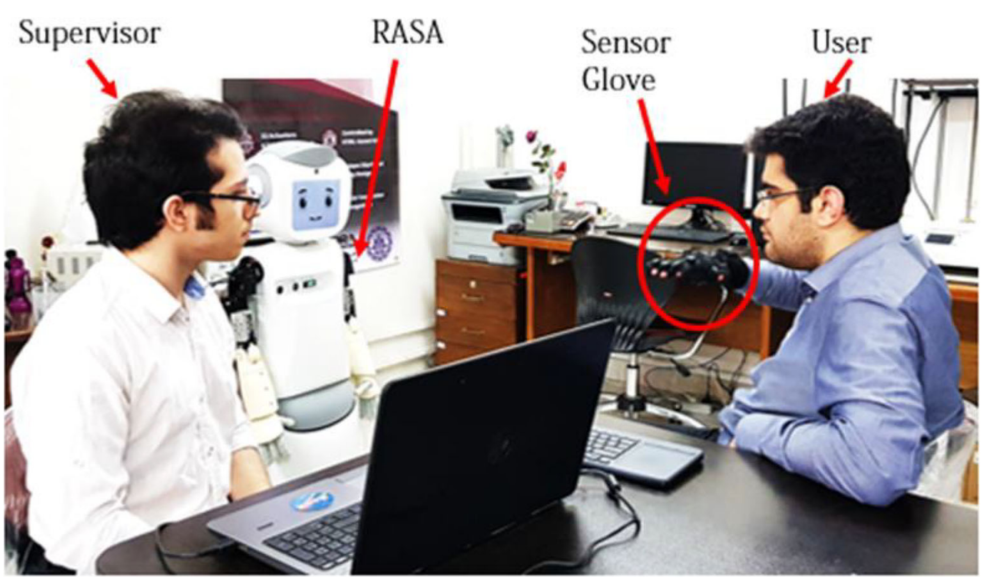

Fig. 7 The experimental setup of the first experiment

Fig. 8 The experimental setup of the second experiment

adapt the level of difficulty in its future training sessions. A summary of the statistical analyzes used for the comparison is shown in Table 1. According to the table, although the average of the second group is less than the first group (i.e., the users in the second group failed on average in fewer words), the *p* value is above 5%. The magnitude of the Cohen's d effect size is reported to be the large value of 1.03 over eight observations, indicating that the difference in the two averages is meaningful and the effects of adaptation in the training sessions are significant. In other words, with the adaptability of the training, the robot was able to identify which sessions it should teach less harshly and which ones more. In future research, this result can be studied and evaluated on more users.

3.2 UTAUT Results: Both Experiments

A summary of the statistical results of the questionnaire items for the first and the second experiment is given in Tables 2 and 3, respectively. In both tables, items with significant differences between the two groups (i.e., p values<0.05 or close to 0.05) are marked in green. As seen, 4 out of the 12 items for the first experiment, and 9 out of 12 items for the second experiment, show significant differences between the two groups. More details of these items are presented in the following subsections.

By comparing the two tables, we can observe that two items in both experiments show a significant difference when we apply the adaptive architecture to the robot. The most important item of the questionnaire in the scope of this study, Perceived Adaptability (PAD), is significantly larger for the second group. This shows that the users could successfully observe the adaptive behavior of RASA, and the proposed methodology of this work has been able to make the teaching program of RASA adaptive. The other item is Anxiety (ANX), again larger for the second group, which will be explained in further detail in the following section.

3.2.1 Anxiety (ANX)

The anxiety of the users is one of the criteria that shows the importance of research on teaching to users. Tables 4 and 5 show the results of this item for the first experiment and second experiment, respectively. In both, the mean of the second group is significantly higher than the first group, and the null hypothesis is rejected. This is one of the more interesting results of this research and shows that the second group of users experienced more average anxiety during the adaptive teaching by the robot. This greater anxiety may have been caused by a greater fear of making mistakes during the teaching. We speculated that since RASA had gained teaching experiences for the second group, and its parameters were well regulated, and the training process was more effective, users

Table 1 The Statistical Analysis Related to the Users' Number of Failed Words in the first experiment

	1st group	2nd group
1st user	6	3
2nd user	5	4
3rd user	3	3
4th user	3	2
Mean	4.25	3
variance	2.25	0.67
number of observations	8	
P(T<=t) one-tail	0.1	
P(T<=t) two-tail	0.2	
Cohen's d	−1.03	

Table 2 A Summary of the Statistical Analysis of the Items of the Questionnaire for the First Experiment

Item	Group 1 mean (var)	Group 2 mean (var)	P Value two-tails	Cohen's d effect size	Number of observations
ANX	1.94 (1.00)	2.75 (1.13)	0.03	0.79	32
ATT	3.58 (1.17)	3.92 (0.99)	0.44	0.32	24
FC	3.5 (0.57)	3.63 (1.13)	0.79	0.14	16
ITU	3 (0.57)	3.25 (0.92)	0.70	0.28	8
PAD	3.5 (1.00)	4.1 (0.52)	0.04	0.69	40
PENJ	3.94 (0.33)	4.13 (0.52)	0.42	0.29	32
PEOU	3.25 (0.50)	3.5 (1.43)	0.62	0.25	16
PS	2.38 (1.41)	2.25 (1.36)	0.83	0.11	16
PU	3.63 (0.27)	4.25 (0.50)	0.06	1.01	16
SI	3.5 (0.33)	4 (0.67)	0.36	0.71	8
SP	2.67 (0.97)	1.67 (0.42)	0.01	1.42	24
TRUST	3.63 (1.13)	3.63 (1.41)	1.00	0	16

felt more pressure in the robotic training. Their point of view toward the process became more serious, making them pay more attention to the answer they give to the robot, and this increased their anxiety.

3.2.2 Perceived Adaptiveness (PAD)

This item is the primary focus of this research, and its results are vital to evaluate the methodology of our study. This item is used to measure how effective the implemented adaptive teaching system has been in this study and to what extent this adaptability was felt by the users. The results of the statistical analysis related to this section are presented in Tables 6 and 7 for the first and the second experiments, respectively. According to the tables, the mean of the second group is significantly larger than the first group, and the null hypothesis is rejected. This means that the more the robot learns and adapts its output parameters to the users and the words, the more the users feel this adaptation, and in practice, the robot was able to better adapt itself over time. In fact, the robot knew the combination of the following occurrences better for the second group: a) which words were harder and which were easier, b) how fast and how many times each of them should be performed, and also c) how it could better express its emotional response, and the results indicate that this was felt by the users. Therefore, we conclude that the methodology used in this research can be a suitable practical solution for adapting the teaching program of social robots. This is in line with the results of [16], where a robot's adaptability is reported to enhance the individual's perception of the system's understanding.

3.3 UTAUT Results: First Experiment

In this section, we will go through the other items that showed a significant difference between the two groups in the first

Table 3 A Summary of the Statistical Analysis of the Items of the Questionnaire for the Second Experiment

Item	Group 1 mean (var)	Group 2 mean (var)	P Value two-tails	Cohen's d effect size	Number of observations
ANX	1.94 (1.18)	2.28 (1.28)	0.03	0.31	192
ATT	3.10 (1.28)	3.53 (1.35)	0.03	0.38	144
FC	2.62 (1.00)	3.50 (0.94)	< 0.01	0.89	96
ITU	3.00 (1.04)	3.29 (1.26)	0.35	0.27	48
PAD	3.06 (0.66)	3.73 (0.62)	<0.01	0.83	240
PENJ	3.49 (1.35)	3.89 (1.00)	0.01	0.36	192
PEOU	2.50 (0.85)	3.46 (1.27)	<0.01	0.93	96
PS	2.15 (1.02)	2.58 (0.84)	0.03	0.45	96
PU	3.81 (0.79)	3.90 (0.95)	0.66	0.09	96
SI	3.38 (0.85)	3.92 (0.60)	0.03	0.64	48
SP	2.24 (0.97)	2.35 (1.22)	0.52	0.11	144
TRUST	2.94 (1.38)	3.63 (1.22)	< 0.01	0.60	96

Table 4 The Statistical Analysis Related to the Anxiety Item for the first experiment

t-Test: Two-Sample Assuming Unequal Variances		
	Group 1	Group 2
Mean	1.94	2.75
Variance	1.00	1.13
Observations	16	16
Hypothesized Mean Difference	0	
df	30.00	
t Stat	−2.23	
P(T<=t) one-tail	0.02	
t Critical one-tail	1.70	
P(T<=t) two-tail	0.03	
t Critical two-tail	2.04	
Total Observations	32	
Cohen's d	0.79	

Table 6 The Statistical Analysis Related to the Perceived Adaptiveness for the first experiment

t-Test: Two-Sample Assuming Unequal Variances		
	Groupe 1	Groupe 2
Mean	3.50	4.10
Variance	1.00	0.52
Observations	20	20
Hypothesized Mean Difference	0	
df	34.00	
t Stat	−2.18	
P (T<=t) one-tail	0.02	
t Critical one-tail	1.69	
P (T<=t) two-tail	0.04	
t Critical two-tail	2.03	
Total Observations	40	
Cohen's d	0.69	

experiment. One important item to be evaluated in this experiment is the perceived usefulness. In other words, how useful the robot's adaptive teaching is for users. The results of the statistical analyzes performed on this item are given in Table 8. According to the table, the mean of the second group in this item is marginally significantly larger than the first group (with a p value of 6% and an effect size of 1.01 on a total of 16 observations), and the null hypothesis is again very close to being rejected. As expected, as the robot's behavior becomes more adaptive and its outputs more consistent with the user and the teaching content, users also had more positive feedback on the usefulness of robotics training. Therefore, it can be concluded that the more adaptive the behavior of a teaching assistant robot, the more useful the robotic training is from the users' point of view. This result is in line with [14, 15], where adaptability is perceived to increase the usefulness of the tutoring system by keeping the child users engaged for a longer period of time. Moreover, in accordance with [14, 15], as both studies we concluded that users perceive artificial tutoring systems as advantageous.

One of the results obtained in the first experiment is related to an item of Social Presence from the questionnaire. The results of the statistical analysis on this item are given in Table 9. Contrary to expectations, the mean of the second

Table 5 The Statistical Analysis Related to the Anxiety Item for the second experiment

t-Test: Two-Sample Assuming Unequal Variances		
	Group 1	Group 2
Mean	1.94	2.28
Variance	1.18	1.28
Observations	96	96
Hypothesized Mean Difference	0	
df	190	
t Stat	−2.15	
P(T<=t) one-tail	0.02	
t Critical one- tail	1.66	
P(T<=t) two-tail	0.03	
t Critical two-tail	1.97	
Total Observations	192	
Cohen's d	0.31	

Table 7 The Statistical Analysis Related to the Perceived Adaptiveness for the second experiment

t-Test: Two- Sample Assuming Unequal Variances		
	Group 1	Group 2
Mean	3.06	3.73
Variance	0.66	0.62
Observations	120	120
Hypothesized Mean Difference	0	
df	238	
t Stat	−6.45	
P(T<=t) one-tail	3.08E-10	
t Critical one-tail	1.65	
P(T<=t) two-tail	6.16E-10	
t Critical two-tail	1.97	
Total Observations	24	
Cohen's d	0.84	

Table 8 The Statistical Analysis Related to the Perceived Usefulness

t-Test: Two- Sample Assuming Unequal Variances

	Groupe 1	Groupe 2
Mean	3.63	4.25
Variance	0.27	0.50
Observations	8	8
Hypothesized Mean Difference	0	
df	13.00	
t Stat	−2.02	
P(T<=t) one-tail	0.03	
t Critical one-tail	1.77	
P(T<=t) two-tail	0.06	
t Critical two-tail	2.16	
Total Observations	16	
Cohen's d	1.01	

group is significantly smaller than the first group (*p* value of 1% and an effect size of 1.2 on 24 observations), rejecting the null hypothesis, which means that the second group felt less that RASA is a real being with real feelings. We suppose that this could be due to the robot's behavior becoming more adaptive, and as the users become more involved in the training process, they monitored the robot's teaching more carefully. Questioning the nature of the robot more than the first group could lead the second group to greater interest and focus on the robot's feelings and behavior, drawing attention to the fact they are not be dealing with a real being. This conjecture could form the basis for an important study in the field of social robotics, this time with more observations and clinical trials to determine its soundness.

Table 9 The Statistical Analysis Related to the Social Presence Item

t-Test: Two- Sample Assuming Unequal Variances

	Groupe 1	Groupe 2
Mean	2.67	1.67
Variance	0.97	0.42
Observations	12	12
Hypothesized Mean Difference	0	
df	19.00	
t Stat	2.93	
P(T<=t) one-tail	0.00	
t Critical one-tail	1.73	
P(T<=t) two-tail	0.01	
t Critical two-tail	2.09	
Total Observations	24	
Cohen's d	−1.20	

3.4 UTAUT Results: Second Experiment

In this section, we will go through all the other items that showed a significant difference between the two groups in the second experiment, but not the first. In the second experiment, 7 out of 12 UTAUT items showed a significant difference between the two groups of participants but showed no significant difference in the first experiment. This is highly likely due to the fact that the number of participants in the second experiment is six times the first, revealing a higher statistical validity.

The participants of the second group showed a significantly more positive attitude toward robotic teaching (ATT), with a *p* value of 0.03 (Table 10). This indicates that adapting the robot's teaching program to the users results in an improvement in the attitude toward the use of robots at teachers. Also, the participants of the second group showed that it would be easier for them to use robotic teaching (FC) when the robot behaves adaptively. This difference is obtained with a p value smaller than 1% (Table 11). This shows the same trend as the PEOU item, meaning that the users in the second group think that the ease of use for the adaptive RASA is higher than non-adaptive RASA (with a p value below 1% and a Cohen's d effect size of 0.93). It is also worth noting that the SI item for the second group is also larger (with a p value of 0.03 and a Cohen's d effect size of 0.64), indicating that the second group participants have, on average, a higher opinion of the future of robotic teaching on their lives. These results are also in-line with the findings of [37], where statistical analysis on different UTAUT items is performed to assess the effects of more and less adaptive versions of an assistive social agent.

The users in the second group showed significantly more interest in the robotic teaching by having larger perceived joy

Table 10 The Statistical Analysis Related to the ATT Item

t-Test: Two- Sample Assuming Unequal Variances

	Group 1	Group 2
Mean	3.1	3.53
Variance	1.28	1.35
Observations	72	72
Hypothesized Mean Difference	0	
df	142	
t Stat	−2.26	
P(T<=t) one-tail	0.01	
t Critical one-tail	1.66	
P(T<=t) two-tail	0.03	
t Critical two-tail	1.98	
Total Observations	144	
Cohen's d	0.38	

Table 11 The Statistical Analysis Related to the FC Item

t-Test: Two- Sample Assuming Unequal Variances

	Group 1	Group 2
Mean	2.62	3.5
Variance	1	0.94
Observations	48	48
Hypothesized Mean Difference	0	
df	94	
t Stat	−4.35	
P(T<=t) one-tail	1.72E-05	
t Critical one-tail	1.66	
P(T<=t) two-tail	3.44E-05	
t Critical two-tail	1.99	
Total Observations	96	
Cohen's d	0.89	

score than the first group. We speculate that this rise in the users' joy (PENJ) is due to the adaptive behavior of the robot, which allows the users to better interact with the robot's dynamic facial expressions and physical representation. The statistical results are shown in Table 12.

Furthermore, the participants of the second group perceived the robot's ability to perform social behavior (PS) as significantly higher. The related statistical analysis is shown in Table 13. We speculate that the meaningful emotional valence variations during teaching for the second group are responsible for this result.

Lastly, the users in the second group scored significantly higher in their trust in RASA and the robotic teaching. The statistical results are shown in Table 14. Our interpretation is that when the behavior and the teaching program of the robot are adaptive, the users more often think that the robot knows what it is doing, and since the teaching it provides better suits them, it leads to having more trust and confidence in the robot.

4 Limitations & Future Work

In addition to examining the effectiveness of the mechanism used in this research, the results of this study are reported as preliminary exploratory findings to create the basis and context of a future research question in the field of social robotics. Moreover, the correctness of these hypotheses should be tested through statistical analysis on a larger target community determined by clinical training in an independent study.

Several technical limitations and difficulties arise while applying this methodology in a real-world scenario. The data-capturing system or the sensory glove can be counted as the first limitation because it is not easy to initialize and use for users during long interactions. Secondly, the start/finish point of the user's performance is determined and fed to the system by a human supervisor, which does not allow the system to run entirely automatically. This is also responsible for some of the Sign Recognition Module faults, meaning the robot may fail a user after a correct performance. Finally, since the fuzzy membership functions are chosen to be linear for the sake of teaching feasibility in real-time scenarios, the robot's initial judgments are too "sharp", meaning the changes in robot's teaching behavior in the initial teaching sessions are too drastic. This problem is resolved after a few teaching sessions when the robot has gathered enough data to change its output parameters in a logical manner.

Table 12 The Statistical Analysis Related to the PENJ Item

t-Test: Two- Sample Assuming Unequal Variances

	Group 1	Group 2
Mean	3.49	3.89
Variance	1.35	1
Observations	96	96
Hypothesized Mean Difference	0	
df	186	
t Stat	−2.53	
P(T<=t) one-tail	0.01	
t Critical one-tail	1.65	
P(T<=t) two-tail	0.01	
t Critical two-tail	1.97	
Total Observations	192	
Cohen's d	0.36	

Table 13 The Statistical Analysis Related to the PS Item

t-Test: Two- Sample Assuming Unequal Variances

	Group 1	Group 2
Mean	2.15	2.58
Variance	1.02	0.84
Observations	48	48
Hypothesized Mean Difference	0	
df	93	
t Stat	−2.22	
P(T<=t) one-tail	0.01	
t Critical one-tail	1.66	
P(T<=t) two-tail	0.03	
t Critical two-tail	1.99	
Total Observations	96	
Cohen's d	0.45	

Table 14 The Statistical Analysis Related to the TRUST Item

t-Test: Two- Sample Assuming Unequal Variances

	Group 1	Group 2
Mean	2.94	3.63
Variance	1.38	1.22
Observations	48	48
Hypothesized Mean Difference	0	
df	94	
t Stat	−2.96	
P(T<=t) one-tail	197E-03	
t Critical one-tail	1.66	
P(T<=t) two-tail	3.94E-03	
t Critical two-tail	1.99	
Total Observations	96	
Cohen's d	0.6	

The parameters of the fuzzy membership functions are not tuned during the training session; It is the user's and the word's parameters that are changed after any training session, which changes the user's/word's score accordingly and thus their position inside the membership functions. Therefore, the adaptation is performed by changing the position of the user/word inside their respective fixed membership functions. For future studies, the authors would suggest implementing an algorithm that modifies the shape of the membership functions while simultaneously changing the scores of the users and the words, and then compare the system's efficiency with this study. The authors would also suggest expanding the vocabulary domain of the robot and teaching individuals on a broader set of words for more comprehensive teaching scenarios.

5 Modeling Errors and Uncertainties

In this research, the term "adaptive" is used to signify that the robot's teaching parameters are changed with respect to the performance of the users during training sessions. The adaptation process is conducted using a fuzzy algorithm that takes the user's and the word's parameters as an input and generates the output by fuzzy rule base and fuzzy database. Therefore, the "learning" module of the algorithm works with the input data that are extracted from users/words profiles, that contain average accuracy, average speed, and average repeat time for each. The average repeat time and the average speed of performance are data that are immune to noises, since they can be easily counted/measured by the robot. The average accuracy is calculated by the deep neural network in the Sign Recognition Module, and it reports the accuracy based on the human supervisor's time framing and the sensory data coming from the data-capturing glove. Although deep neural networks have shown to be more robust against a certain amount of noises compared to other regression/classification methods, there are noises in the sensory data of the glove and uncertainties in the human supervisor time framing or robot's motor function data that might either misdirect the user to perform the incorrect word, or lead to wrong classification by the network; In this case, the robot's judgment of the user would be wrong and consequently the parameters of the particular word and user would be adjusted incorrectly, leading to inefficient teaching parameters for the next teaching sessions. This case is rare however, and since for every training session the robot considers the whole training history of the user and the word and adjust its parameters on the entire data, a small number of faulty adjustments does not affect the teaching process significantly. Therefore, modeling these small uncertainties in our adaptive teaching algorithm would not be either efficient or necessary, and it is not the focus of this research.

Table 15 Items investigated in the standard UTAUT questionnaire

Evoking anxious or emotional reactions when it comes to using the system	ANX
Positive or negative feelings about towards the appliance of the technology	ATT
Factors in the environment that facilitate use of the system	FC
The intention to use the system over a longer period in time	ITU
The perceived ability of the system to adapt to the needs of the user	PAD
The perceived feelings of joy/pleasure associated with the use of the system	PEN J
The degree to which one believes that using the system would be free of effort	PEOU
The perceived ability of the system to perform sociable behavior	PS
The degree to which a person believes that the system would be assistive	PU
The persons perception that people who are important to him think he should or should not use the system	SI
The experience of sensing a social entity when interacting with the system	SP
The belief that the system performs with personal integrity and reliability	TRUST

Table 16 The questions in our proposed UTAUT-based questionnaire

	I find the robotic teaching intimidating	I find the robot scary	During the training, i would be afraid to break something	During the training, I would be afraid to make a mistake	ANX
		During the training, I felt that I'm really learning something	I don't think robots can teach like humans	I think that robot teaching is a good idea	ATT
			I know enough to make use of robot teaching	I have everything I need to use robotic teaching	FC
				In future, I will use rob ots as teachers	ITU
I felt that the robot knew which words are easy or hard	I felt that the robot knew in which words I'm weak or strong	I think the robots gain experience and teaches better over time	I think the robot better adapts to me over time	I think the robot is adaptive to me	PAD
	I enjoyed the adaptation of the robot	I find robotic teaching fascinating	I enjoy interacting with robots	I enjoy robotic Teaching	PENJ
			I know quickly how to use robotic teaching	I find robotic teaching easy	PEOU
			I find the robot a pleasant social partner	I think the robot understands me	PS
			I think it would be convenient to use robots as teachers	I find robotic teaching useful	PU
				It would give a good time impression if I would use robotic teaching	SI
		Sometimes the robot seems to have real feelings	I can imagine the robot to be a living creature	When interacting with robot I felt like it's a real person	SP
			I would follow the advice the robot gives me	I would trust robotic teaching	TRUST

6 Conclusion

Using social robots as a sign language teaching assistant can be an important step in expanding communication with the deaf in the future. The literature shows that user interfaces with adaptive behavior lead to greater user acceptance and increased teaching efficiency compared to non-adaptive ones. This project aims to empower the RASA social robot for adaptive teaching of Iranian Sign Language to users, which first requires the robot's logic to be adjusted toward the teaching content in the next training sessions, and secondly, adapt its methods to intelligently teach based on the users past and present performance. By designing and implementing an architecture to produce adaptive parameters appropriate to both the user and the content of the training, and then have the robot perform sign language words according to these parameters, we were able to enable the robot to adaptively teach Iranian Sign Language in a manner somewhat similar to a human teacher. After the general mechanism of the architecture was explained, each of the modules that make up this architecture was discussed, and their performance was explained.

In two separate experiments, users (8 in-person and 48 virtually) with different levels of familiarity with robotics and Iranian Sign Language were recruited and divided into two groups. Each user was taught for ten minutes. In the first experiment, the users wore robotic data-collection gloves in front of the robot, and their training process was monitored by a supervisor. In the second experiment, the users were shown a first-person view video of the robot performing SL teaching, adaptive and nonadaptive. In both experiments, the adaptive teaching architecture was applied only for the second group. Thus, the effect of the adaptive teaching was measured by comparing the average parameters of users in the two groups in terms of the number the words they failed and also with various items from the standard questionnaire UTAUT from the social robotics field. Utilizing the results of this questionnaire, the statistical analysis of the T-test and the Cohen's d effect size were performed to compare the mean of the two groups in different items. Initially, it was found that the second group of users failed significantly fewer words than the first group, which shows that adapting the robot causes it to teach more intelligently and effectively, as well as how the robot was able to adjust the teaching parameters toward different users.

Two of the 12 UTAUT items showed significant differences between the two groups in both experiments, according to the T-test results and effect size tests. One of these items is anxiety, which indicates that the second group felt moderately more anxious than the first group. Moreover, the second group felt the robot's adaptability more than the first group (both groups believed that the robot's teaching was adaptive), which could indicate that the methodology used in this study was effective and that the robot's compatibility was felt by users. As previously explained in the results section, two other items in the first experiment (PU and SP) and seven other items in the second experiment (ATT, FC, PEOU, PENJ, PS, SI, and TRUST) showed a significant difference between the two groups. The general mechanism of this methodology is also transferable to other social robots and other teaching scenarios, as no specific characteristics of RASA or Iranian Sign Language were engaged in the architecture [38–40].

Acknowledgments This research was supported by the Iranian National Science Foundation (INSF) (http://en.insf.org/). The complementary and continues support of the Social & Cognitive Robotics Laboratory by Dr. Ali Akbar Siassi Memorial Grant is also greatly appreciated.

Availability of Data and Material (Data Transparency) All data from this project (videos of the sessions, results of the questionnaires, scores of performances, etc.) are available in the archive of the Social & Cognitive Robotics Laboratory.

Code Availability All of the codes are available in the archive of the Social & Cognitive Robotics Laboratory. In case the readers need the codes, they may contact the corresponding author.

Authors' Contributions All authors contributed to the study conception and design, material preparation, data collection and analysis were performed by Salar Basiri and Alireza Taheri. The first draft of the manuscript was written by Salar Basiri and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research was funded by the Iranian National Science Foundation (INSF) (http://en.insf.org/) (Grant No. 98025100)

Declarations

Conflicts of Interest/Competing Interests Author Alireza Taheri has received research grants from the Iranian National Science Foundation (INSF). The authors Salar Basiri, Ali Meghdari, and Minoo Alemi declare that they have no conflict of interest.

Ethics Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Ethical approval for the protocol of this study was provided by the Iran University of Medical Sciences (#IR.IUMS.REC.1395.95301469).

Consent to Participate Informed consent was obtained from all individual participants included in the study.

Consent for Publication The authors affirm that human research participants provided informed consent for publication of the image in Fig. 7. All of the participants have consented to the submission of the results of this study to the journal.

References

1. Martins, G.S., Santos, L., Dias, J.: User-adaptive interaction in social robots: a survey focusing on non-physical interaction. Int. J. Soc. Robot. **11**(1), 185–205 (2018)
2. Zakipour, M., Meghdari, A., Alem, M.: RASA: a Low-Cost Upper-Torso Social Robot Acting as a Sign Language Teaching Assistant, " Presented at the International Conference on Social Robotics (2016)
3. C. Courtin, "The impact of sign language on the cognitive development of deaf children: the case of theories of mind," J. Deaf. Stud. Deaf. Educ., vol. 5, no. 3, pp. 266–276, Summer 2000
4. Norcio, A.F., Stanley, J.: Adaptive human-computer interfaces: a literature survey and perspective. IEEE Trans. Syst. Man. Cybernetics. **19**(2), 399–408 (1989)
5. McTear, M.F.: User modelling for adaptive computer systems: a survey of recent developments. Artif. Intell. Rev. **7**(3–4), 157–184 (1993)
6. Rossi, S., Ferland, F., Tapus, A.: User profiling and behavioral adaptation for HRI: a survey. Pattern Recogn. Lett. **99**, 3–12 (2017)
7. Karami, A.B., Mouaddib, A.I.: A decision model of adaptive interaction selection for a robot companion, pp. 83–88. ECMR (2011) Corpus ID: 15367459
8. Karami, A.-B., Sehaba, K., Encelle, B.: Adaptive and personalised robots - learning from users' feedback. In: IEEE 25th International Conference on Tools with Artificial Intelligence, Washington DC, United States (Nov 2013). https://doi.org/10.1109/ICTAI.2013.98
9. Karami, A.-B., Sehaba, K., Encelle, B.: Adaptive artificial companions learning from users' feedback. Adaptive Behavior, SAGE Publications. **24**(2), 69–86 (2016). https://doi.org/10.1177/1059712316634062〉
10. Grosinger. J., Pecora, F., Saffiotti, A.: Making robots proactive through equilibrium maintenance. In: 25th International Joint Conference on Artificial Intelligence [Internet]. 2016. Available from: http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-51100 (March 2021)
11. Amirabdollahian, F., et al.: Accompany: Acceptable roboticS COMPanions for AgeiNG Years — Multidimensional aspects of human-system interactions. In: 2013 6th International Conference on Human System Interactions (HSI), pp. 570–577 (2013). https://doi.org/10.1109/HSI.2013.6577882
12. Aly, A., Tapus, A.: Towards enhancing human-robot relationship: customized robot's behavior to human's profile. In: The AAAI Fall Symposium Series (AI-HRI), Arlington, Virginia, United States (Nov 2014) 〈(hal-01265965〉
13. Westlunda, J.K., Gordona, G., Spauldinga, S., Leea, J.J., Plummera, L., Martinezb, M., Dasb, M., Breazeal, C.: Learning a second language with a socially assistive robot. Corpus ID. 14461956 (2015)
14. Gordon, G., Breazeal, C.: Bayesian Active Learning-Based Robot Tutor for Children's Word-Reading Skills," Presented at the National Conference on Artificial Intelligence (2015)
15. Gordon, G., Spaulding, S., Westlund, J.K., Lee, J.J., Plummer, L., Martinez, M., Das, M., Breazeal, C.: Affective personalization of a social robot tutor for children's second language skills. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI'16), pp. 3951–3957. AAAI Press (2016)
16. Aylett, R., et al.: I Know how that Feels – an Empathic Robot Tutor, in *eChallenges e-2015 Conference Proceedings*. Vilnius, Lithuania (2015)
17. Tozadore, D.: Project r-Castle: Robotic-Cognitive Adaptive System for Teaching and Learning (2017)
18. Almeida, O., de Magalhaes Netto, J.F.: Adaptive educational resource model to promote robotic teaching in STEM Courses. In:

2019 IEEE Frontiers in Education Conference (FIE), pp. 1–8 (2019). https://doi.org/10.1109/FIE43999.2019.9028417
19. Li, X., O'Regan, D., Akca, H.: Global exponential stabilization of impulsive neural networks with unbounded continuously distributed delays. IMA J. Appl. Math. **80**(1), 85–99 (2013)
20. Zhang, X., He, S., Stojanovic, V., et al.: Finite-time asynchronous dissipative filtering of conic-type nonlinear Markov jump systems. Sci. China Inf. Sci. **64**, 152206 (2021). https://doi.org/10.1007/s11432-020-2913-x
21. Stojanovic, V., He, S., Zhang, B.: State and parameter joint estimation of linear stochastic systems in presence of faults and non-Gaussian noises. Intl. J. Robust Nonlinear Control. **30**(16), 6683–6700 (2020)
22. Zhou, L., Tao, H., Paszke, W., Stojanovic, V., Yang, H.: PD-Type Iterative Learning Control for Uncertain Spatially Interconnected Systems," Mathematics*,* vol. 8, no. 9, (2020)
23. Chen, Z., Zhang, B., Stojanovic, V., Zhang, Y., Zhang, Z.: Event-based fuzzy control for T-S fuzzy networked systems with various data missing. Neurocomputing. **417**, 322–332 (2020)
24. Zadeh, L.A.: Fuzzy sets. Inf. Control. **8**(3), 338–353 (1965)
25. Carinena, P., Regueiro, C. V., Otero, A., Bugarin, A. J., Barro, S.: Landmark detection in mobile robotics using fuzzy temporal rules," vol. 12, no. 4%J Trans. Fuz Sys., pp. 423–435 (2004)
26. Huq, R., Mann, G. K. I., Gosine, R. G.: Behavior-modulation technique in mobile robotics using fuzzy discrete event system. IEEE Transactions on Robotics*,* vol. 22, no. 5, (2006)
27. Shaukat, A., Burroughes, G., Gao, Y.: Self-Reconfigurable Robotics Architecture Utilising Fuzzy and Deliberative Reasoning," Presented at the 2015 SAI Intelligent Systems Conference (IntelliSys), London, UK (2015)
28. Fateh, M.M., Khorashadizadeh, S.: Robust control of electrically driven robots by adaptive fuzzy estimation of uncertainty. Nonlinear Dynamics. **69**(3), 1465–1477 (2012)
29. Wu, C., Liu, J., Jing, X., Li, H., Wu, L.: Adaptive fuzzy control for nonlinear networked control systems. IEEE Trans. Syst. Man Cybernet.: Syst. **47**(8), 2420–2430 (2017)
30. Pelletier, F.J.: Petr Hájek. Metamathematics of fuzzy logic. Trends in logic, vol. 4. Kluwer academic publishers, Dordrecht, 1998, viii + 297 pp," Bull Symbolic Logic*,* vol. 6, no. 3, pp. 342–346, 2014
31. Zadeh, L.A.: Fuzzy sets. Inf Control. **8**(3), 338–353 (1965). https://doi.org/10.1142/9789814261302_0021
32. Belpaeme, T., Vogt, P., van den Berghe, R., Bergmann, K., Göksun, T., de Haas, M., Kanero, J., Kennedy, J., Küntay, A.C., Oudgenoeg-Paz, O., Papadopoulos, F., Schodde, T., Verhagen, J., Wallbridge, C.D., Willemsen, B., de Wit, J., Geçkin, V., Hoffmann, L., Kopp, S., Krahmer, E., Mamus, E., Montanier, J.M., Oranç, C., Pandey, A.K.: Guidelines for designing social robots as second language tutors. Int. J. Soc. Robot. **10**(3), 325–341 (2018)
33. Paléologue, V., Martin, J., Pandey, A.K., Coninx, A., Chetouani, M.: Semantic-based interaction for teaching robot behavior compositions. In: 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 50–55 (2017). https://doi.org/10.1109/ROMAN.2017.8172279
34. Meghdari, A., Alemi, M., Zakipour, M., Kashanian, S.A.: Design and realization of a sign language educational humanoid robot. J. Intell. Robotic. Syst. **95**(1), 3–17 (2018)
35. Krägeloh, C.U., Bharatharaj, J., Sasthan Kutty, S.K., Nirmala, P.R., Huang, L.: Questionnaires to measure acceptability of social robots: a critical review. Robotics. **8**, 88 (2019). https://doi.org/10.3390/robotics8040088
36. Williams, M.D., Rana, N.P., Dwivedi, Y.K.: The unified theory of acceptance and use of technology (UTAUT): a literature review. J. Enterp. Inf. Manag. **28**(3), 443–488 (2015)
37. Heerink, M., Kröse, B., Evers, V., Wielinga, B.: Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int. J. Soc. Robot. **2**(4), 361–375 (2010)
38. Meghdari, A., Alemi, M.: STEM teaching-learning communication strategies for deaf students, Proc. 17th international RAIS conference on social sciences and humanities, pp. 47–55. Princeton University, USA (2020)
39. Alemi, M., Meghdari, A.: Science and engineering education strategies along with the deaf-sign language. Iranian J. Eng. Educ. **22**(85), 69–84, Spring (2020)
40. Alemi, M., Taheri, A., Shariati, A., Meghdari, A.: Social Robotics, Education, and Religion in the Islamic World: An Iranian Perspective. J. Sci. Eng. Ethics, vol. **26**, no. 3, (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Salar Basiri received his B.S. degree in Mechanical Engineering from Sharif University of Technology (SUT), Tehran, Iran in 2018. He received his M.S. degree in Mechatronics from Sharif University of Technology (SUT), Tehran, Iran in 2020. His research interests include Mechatronics, Artificial Intelligence, and Machine Learning.

Alireza Taheri is an Assistant Professor of Mechanical Engineering with an emphasis on Social and Cognitive Robotics at Sharif University of Technology, Tehran, Iran. He is a Research Associate and the Technical Administrator at the Social and Cognitive Robotics Lab at Sharif University of Technology.

Ali Meghdari is a Professor of Mechanical Engineering and Robotics at Sharif University of Technology (SUT) in Tehran. Professor Meghdari has performed extensive research in various areas of robotics; social and cognitive robotics, mechatronics, bio-robotics, and modeling of biomechanical systems. He has been the recipient of various scholarships and awards, the latest being: the 2012 Allameh Tabatabaei distinguished professorship award by the National Elites Foundation of Iran (BMN), the 2001 Mechanical Engineering Distinguished Professorship Award from the Ministry of Science, Research & Technology (MSRT) in Iran, and the 1997 ISESCO Award in Technology from Morocco. He is currently an affiliate member of the Iranian Academy of Sciences (IAS), a Fellow of the American Society of Mechanical Engineers (ASME), and the Founder and Chancellor of Islamic Azad University-Fereshtegaan International Branch (for students with special needs; primarily the Deaf and Blind, https://fib.iau.ir/en).

Minoo Alemi received her Ph.D. in Applied Linguistics from Allameh Tabataba'i University in 2011. She is currently an Associate Professor and Division Head of Applied Linguistics at the Islamic Azad University, West-Tehran Branch. She is the co-founder of Social Robotics in Iran, a title she achieved as a Post-Doctoral research associate at the Social Robotics Laboratory of the Sharif University of Technology. Her areas of interest include discourse analysis, interlanguage pragmatics, materials development, and RALL. Dr. Alemi has been the recipient of various teaching and research awards from Sharif University of Technology, Allameh Tabataba'i University, Islamic Azad University, and Int. Conf. on Social Robotics (ICSR-2014)