Review Article
Volume 10:7, 2021

Advances in Robotics and Automation

ISSN: 2168-9695 Open Access

An Artificial Intelligence Design Principle Toolkit for Teaching Basic Mathematics to Children

Fatmah Abdulrahman Baothman*

Department of Computing and Information Technology, King Abdul Aziz University, Jeddah, Saudi Arabia

Abstract

There is a constant change in the techniques of teaching and learning as a result of artificial Intelligence (AI). In the past, the objective was to provide an intelligent tutoring system without intervention from a human teacher to enhance skills, control, knowledge construction, with emotional and intellectual engagement. The present study focuses on enhancing the learning capabilities of humanoid robot Nao and increasing his intelligence by activating multi-sensory perceptions among 3 to 7 years old children. Here, the multi-sensory perceptions include; visual and auditory stimuli modules, speech-related, and body movements. The design and testing processes were conducted by implementing an AI principle design, namely the three-constituent principle. The study has developed a toolkit that uses a mixed reality system architecture featuring different ways of interaction between a child and a robot Nao agent. The toolkit enables Arabic speech recognition, and the Haar algorithm was implemented for robust image recognition to improve the capabilities of the Nao during interactions with a child in a mixed reality system.

Keywords: Artificial intelligence • Assistive technologies • Emergent behavior • Mixed reality • Nao humanoid robo

Introduction

Artificial Intelligence (AI) was introduced half a century ago. Researchers initially wanted to build an electronic brain equipped with a natural form of intelligence. The concept of AI was heralded by Alan Turing in 1950s, who proposed the Turing test to measure a form of natural language (symbolic) communication between humans and machines. In the 1960s, Lutfi Zadah proposed fuzzy logic with dominant features of knowledge representation and mobile robots. Stanford University created the Automated Mathematician to explore new mathematical theories based on a heuristic algorithm; although, AI had become unpopular in the 1970s due to its inability to meet unrealistic expectations. The 1980s offered promise for AI as sales of AI-based hardware and software for decision support applications exceeded $400 million. By the 1990s, AI had entered a new era by integrating Intelligent Agent (IA) applications into different fields, such as games, controlling spacecraft, security (credit card fraud detection, face recognition), and transportation (automated scheduling systems). The beginning of the 21st century witnessed significant advances in AI in industrial business and government services with several initiatives, such as intelligent cities, intelligent economy, intelligent industry, and intelligent robots [1].

A unified definition of AI has not yet offered; however, its concept can be built from different definitions;

- It is an interdisciplinary science because it interacts with cognitive science.
- It uses creative techniques in modeling and mapping to improve average performance when solving complex problems.
- It implements different processes to imitate intelligent human or animal behavior. Fourth, the developed system is either a virtual or a physical system with intelligent characteristics.
- It attempts to duplicate human mental and sensory systems to model aspects of the thoughts and behaviors of humans.
- It passes the intelligence test if it interacts completely with other systems or creatures in the world and in realtime.
- It follows a defined cycle of sense-plan-act.

Considering the above-mentioned definitions, the present study proposes the definition of AI as; "Artificial Intelligence is an interdisciplinary science suitable for implementation in any domain that uses heuristic techniques, modeling, and AI-based design principles to solve complex problems. Single or combined processes of perceiving, reasoning, learning, understanding, and acting jointly can improve system behavior and decision making." The goal of AI is to enable virtual and physical intelligent agents, including humans and/or systems that continuously upgrade their intelligence to attain superintelligence. Agents should be able to fully integrate with one another in learning, teaching, adapting themselves to dynamic environments, communicating logically, and functioning efficiently

***Address to correspondence:** Fatmah Abdulrahman Baothman, Department of Computing and Information Technology, King Abdul Aziz University, Jeddah, Saudi Arabia; E-mail: fbaothman@Kau.edu.sa*

Received: 02 August, 2021; **Accepted:** 16 August, 2021; **Published:** 23 August, 2021

with one another or with other creatures in the world and realtime through sense-plan-act-react cycles [2].

The proposed definition enhances research at the level of experimental design in using multi-sensory technologies to improve the interaction and growth of intelligence by applying the AI design principle. Enhanced interaction between humans and robots improves learning, especially in the case of a child. Motion and speech sensor nodes are fused to this end. Children nowadays are familiar with the use of handheld devices like mobile phones, tablets, pads, and Virtual Reality cameras. Therefore, the toolkit developed in this study uses a mixed reality system featuring different ways of interaction between a child and a robot agent. This study makes the following contributions [3].

The three-constituents principle

The three-constituents principle for an agent suggests that "designing an intelligent agent involves constituents; definition of the ecological niche, definition of the desired behaviors and tasks, and design of the agent". Therefore, an agent's intelligence is enabled to grow in time using "here and now," perspective during interactions in different dynamic environments. In the present study, the design of the robotic agent Nao is not among the required tasks, but the other two constituents are related to the educational environment and involve interactions with a human agent. Therefore, this study defines the ecological niche using different environments (a classroom, a lab, and a home) with a focus on a mixed-reality system in which Nao functions are present according to the desired behavior as mathematics is taught to a child. This improves Nao's learning ability and increases his intelligence. Thus, this study shows that constituents; definition of the ecological niche and definition of the desired behaviors and tasks are sufficient to increase Nao's intelligence.

- The "here and now," perspective is related to three-time frames and shows that the behavior of any agent's system matures over a period of time and is associated with three states;
- State-oriented - describes the actual mechanism of the agent at any instance of time.
- Learning and development - relates to learning and development from state-oriented action.
- Evolutionary - the phylogenetic perspective which explains the emergence of a higher level of cognition by emphasizing the power of artificial evolution and performing more complex tasks.

The agent matures into an "adult" by which the process in any state is affected by its previous state. In this study, the present study has focused on state-oriented and learning and development state to observe its outcome in association with the evolutionary state.

Related Work

Mixed reality system

TouchMe provides a third-person camera view of the system, instead of the first-person view available from the robot. The use of a third-person view is more effective for inexperienced users to interact with the robot. To achieve this, they used a fixed-mobile projector. Socially-aware Interactive Playgrounds use various actuators to provide feedback to children. These actuators include projectors, speakers, and lights. These interactive playgrounds can be placed at different locations, such as schools, streets, and gyms. Humans produce, interpret, and detect social signals (a communicative or informative signal conveyed directly or indirectly). Thus, their social signals can be used to enhance interactions with others [4].

Learning through robots

Various studies have been conducted on teaching humans using robots in various environmental settings. Robot stage implements learning among junior high school students through mixed reality systems. Its creators compared the use of physical and virtual characters in a learning environment. RoboStageenables interactions in robots using voice and physical objects to achieve three stages of events that include; learning, situatedness, and blended. These events help students learn and practice activities, understand an environment, and execute an event. GENTORO uses a robot and a handheld projector to interact with children and perform a storytelling activity. Its creators studied the impact of using a small handheld projector on the storytelling process. They also discussed the effects of using audio interactions instead of text, and the use of a wide-angle lens.

Experimental Design

The experiment initiated at king Abdul-Aziz University with Aldebaran representative was related to a three years old robot Nao, who was not able to speak Arabic nor solve simple mathematics. The study analysis initiated by selecting the suitable artificial intelligence principle design for the study. The goals and tasks of the experiment were defined precisely to increase Nao's intelligence to at least seven years old. The Nao mathematics intelligence measurements were based on solving 100 children exercises of basic addition, subtraction, and multiplication problems with the help of human agents. While Nao reached the level of understanding simple sentences for Arabic language speech recognition. The experiment time scale was set for a total of two years. The aim of the study was to involve the robot Nao in the learning-teaching process using interaction and multi-sensory perceptions by exposing Nao to different environments, agents. However, the present work just focused on the mixed reality environment [5].

NAOqi supports C^{++}, Python and JavaScript programming languages to be used on the robot. There are several built-in modules that include auditory, vision, and recognition. The Nao agent, after recognizing the number by a human agent store it in a database to improve its learning capabilities. Three functions were added to the robot system namely approximation function, objective function, and fitness function to improve Nao learning experience. The main task of the experiment at this stage was to enable Nao to interact with children and answer simple questions in mathematics using hand gestures, speech in Arabic, or both. The children would interact with Nao after activating its vision module to recognize the number of fingers.

In the second experiment, the auditory module was activated and a learning auditory guessing game was used. In this game, a child was asked to calculate the product of two numbers, and the robot reacted by making a clapping sound if the mathematics answer was correct or an alarm if it was not. Next, both vision and auditory modules were activated and were played interchangeably. This teaching game continues until the child learns from earlier errors; while, the Nao recognition system improves continuously as it teaches more children and acquires more data. Thus, an agent would navigate and recall the learned action, when an agent is interacting with another human agent in the same environment. environment at the University laboratory to solve basic math problems. Third, Nao was taken home to spent time observing the family members' behaviors by interacting with them. The author implemented the three-constituents principle to focus on the desired interactive behaviors of Nao within the morphological capability and limitations.

The “here and now” perspective explains actions taken by the Nao such as; answering math questions when asked. The learning and development behavior of Nao is described while answering new questions from previous responses, and evolution is concerned with how Nao’s learning behavior evolved by answering unlearned questions or new behavioral mechanisms emerge. In the here-and-now perspective, the mechanisms and principles concern directly on how behavior comes or how an individual behavior resulting from an agent’s interaction with the environments as explained by Hafner and Moller. The three-time-scale “here and now” led to instant action by Nao corresponding to specific learning–teaching situations and unexpected behavior that enable an emergent un-programmed action. Thus, Nao exhibited a defined behavioral relationship with each specified location.

The sensors and controllers developed their perceptual cues and representation models for each environment, and Nao linked certain tasks to each environment using the available module.. This indicates the existence of the notion of emergence, although this is not the focus of this study. For example, when Nao was placed in the home environment observing Muslim prayers, a desired un-programmed behavior emerged to allow Nao to simulate the human body performing the actions involved in the Muslim prayer, such as bending and hand movements. This shows that agent Nao behaviors can emerge from interactions with the environment. In the final stage of the experiment, the learning was based on mixed reality system environments. Performing teaching-learning tasks over an extended period in different environments caused Nao to become more intelligent and develop a new-task in addition to the desired tasks. The robot was able to recognize additional gestures when placed in a mixed reality environment.

The mixed reality system environment

The mixed reality systems are used in various fields, like Earth science, Engineering, and Medicine. They introduce extensive interaction along with a virtual environment to research. These criteria lead to enhanced interaction from the participant (child). Physical agents in the form of a humanoid robot Nao, with enhanced interaction using a mixed reality system, make the learning process more interesting and effective, in comparison to the use of a virtual agent character. The interaction between humans and robot agents has been investigated for many years. The use of hardware-based devices and computer vision-related techniques for interaction has also been studied. The choice between hardware and software-based solutions leads to a tradeoff between accuracy and ease of interaction. In the case of computer vision-based solutions, marker-based and marker-less techniques are used. The use of gloves in marker-based computer vision techniques provides high accuracy but uses a virtual environment. Haar classifiers, convexity defects, and various other techniques have been implemented as part of a marker-less computer vision approach.

In this study, a camera was used to focus on the working space featuring the robot and the child. Information from the robot was provided to the child using its motions and projections on a screen. The camera was mounted above the area to provide a top-down view. The area of focus was fixed, and interaction between the robot and the child only occurred with this specified locality. This limitation did not impair the system because its main objective was to enable the robot to teach the child, who is in its vicinity in any case. The recognition of the gesture from the child intended for the robot was extracted using various known techniques that affected the research for two reasons; the presence of the hand in front of the face and the lamination of the background. The author used both recognition and handheld device-based interaction between the robot and the child to provide haptic feedback from the child to the system. The robot taught the child how to perform mathematical tasks by projecting pre-recorded audio or video, or gestures. The projection required dimming the lights; whereas, the camera that focused on the robot and the child needed proper illumination. Therefore, appropriate lighting or dimming, or a handheld projector was necessary.

System Architecture

The proposed system is made up of five components: a projector, camera, child, robot, and server (Figure 1).

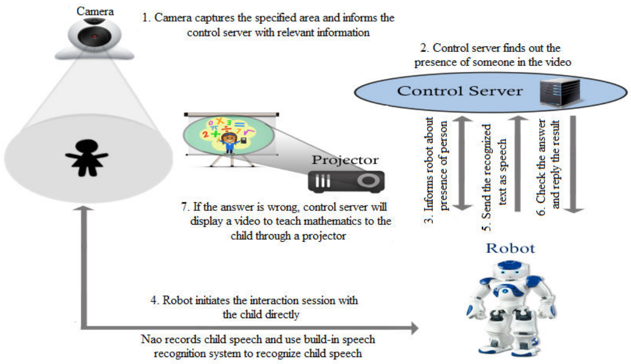

Figure 1. System architecture.

- It focuses on the area of coverage. Its position and focus are fixed. The output of the camera is regularly passed to the server that uses Haar classifiers to identify the child using face recognition. In the absence of the camera, the robot’s eyes are used to input the video, and the face recognition algorithm of the humanoid robot is used to detect the presence of the child.
- It is the main component of the system; whose learning is the sole objective. To make learning fun for the child, he/she is given a variety of choices to choose from, including interaction through a handheld device, speech, or fingers of the hand.
- It is used to detect various finger movements of the child and recognize speech from him/her. It can perform face and speech recognition using its default module or pass the acquired data to the server for processing. A communication server runs on the robot to interact with the handheld device of the child.

- It is used to control the flow of operations. Instructions from the server can be given either to the robot or the projector. The server can also perform face recognition using a Haar classifier or other means. It can also process the detection of the fingers of the child using the convex hull approach and can perform speech recognition.
- It is used to display appropriate learning material for the child. This is done only when there is a need to teach the child through already recorded videos. The process can be affected by light in the coverage area; therefore, a suitable projector has to be used.

Details of toolkit

The proposed system is shown in Figure 2; whereas, provided more detail about the proposed system. The fixed camera and/or Nao'seye is used to acquire the video of the child.

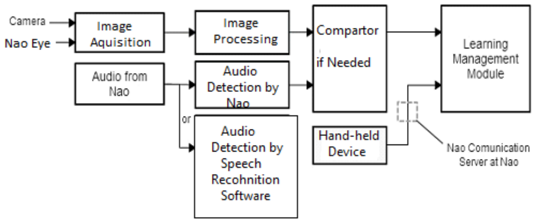

Figure 2. Toolkit description.

The initial image processing is carried out to check whether the child is in front of the robot. The image acquisition and image processing are described in the flowchart. The steps involved in image processing are as follows;

Face identification: The human face in front of the robot is detected using the ALFaceDetection module in Nao. The detected face is written to the ALMemory periodically. Once the Nao robot detects the child's face, it welcomes the child and sends a request to the camera or itself to start acquiring the image.

Restrictive face recognition: Face recognition was regulated by Haarclassifiers on the control server to enable our system to work with only specific children. The data acquired were passed to the control server to recognize already known faces.

Acquisition: Vision was implemented through Nao's eyes. This acquisition was in the form of an image, and a sequence of images on a periodic basis was considered equivalent to a video.

Obtaining image from Nao: The specifications of the image are entered using ALProxy with the ALVideoDevice module. The Generic Video Module (GVM) is used to provide the necessary image format and specifications. The image is obtained using the getImageRemotemethod, and is converted from a pixel image into a PIL image.

Conversion: The obtained image is converted into a grayscale or HSV scale. The author used grayscale, where there are no morphological effects. The HSV scale should be used in case of morphological effects.

Children wearing gloves are processed using HS: Colored gloves are extracted from the image. This separates the fingers of interest from other parts of the acquired image/video.

Morphological effects: Certain morphological effects, like erosion, dilation, and gradient are applied to the image/video in case the acquired image does not meet the requirements of the user. The aim of this process is to enhance the quality of the acquired information.

Blurring: A blur using a Gaussian filter is applied to the gray image. This process is performed to remove Gaussian noise from the image.

Thresholding: The blurred image is then processed in a threshold. This process converts the grayscale image into a binary image. Levels are assigned to pixels based on the threshold value.

Finding contours: Contours refer to the outline of the given image. The hierarchy or relationship between contours is obtained, and they are compressed to save space.

Contour areas: The area of each contour is obtained. The one with the maximum area is identified and passed to the next stage.

System operation

The fixed camera is used to track the position and location of the child. The robot initiates the teaching of simple mathematics, based on the perspective of the child. The child's learning is tested by processing on the server and is based on the output of the learning process. The outcome is displayed on the projected screen by the robot. Thus, both the robot and the projector screen respond to the child. The child also had a mobile device to interact with the robot, and its process of operations are as follows;

- The location of the child is recognized by the camera.
- The child initiates communication with the robot using the mobile application available on his/her mobile device.
- The robot teaches a lesson based on audio, video, and movements. Note that as Naohas only three fingers, so it is not feasible for it to use them to indicate numbers.
- The robot generates a question for the child to answer.
- The child's response is stored and tested using speech recognition tools to check the answer.
- Based on the results of speech recognition, an appropriate response is displayed on the projector by the robot.

Handheld interactions

The system retains ambiguities as despite the image processing. Changes to the background and luminosity impact the quality of image processing, which affects finger recognition. To handle this, a handheld (android) device-based application was used that communicates with the communication server in Nao. The communication server waits for a connection at its socket. The user interface of the android application contains numbers from zero to nine as well as a few mathematical operations (Figure 3). The request is passed to the communication server in Nao, when the child clicks on the operands, operation, and result. It checks the result. If it is correct, it greets the child with a sound and some action. If not, a video on the specific operation is projected for the child from the control server. Options are provided for the child or a parent to change the volume and the speed of the voice generated by the robot.

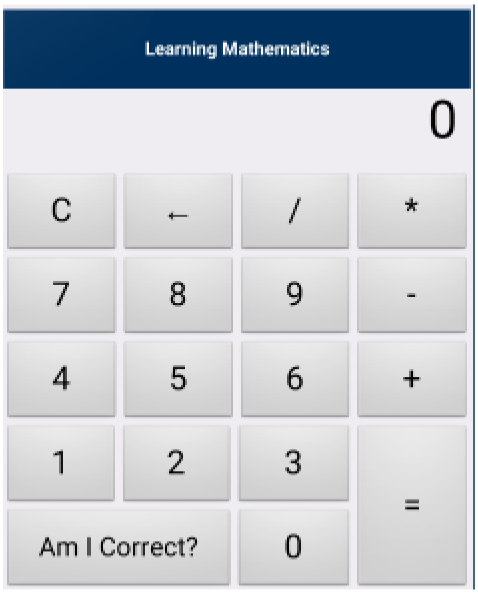

Figure 3. The initial screen of the handheld application.

Results and Discussion

The mixed reality system uses different entities to enhance learning. The performance of the system when using the Haar classifier and Nao'sfacial recognition algorithm was compared. Both correctly identified the face in front of the robot. The drawback of Nao's face recognition was that it expected the face to be close and took some time to recognize it. According to documentation for the ALVisionRecognition-Nao Software 1.14.5 the recognition process works between half and twice the distance used for learning purposes. The closeness of the human face to Naorestricted the free movements of the child is indicated in Table 1.

Criteria	Measure (m)
Communication distance between Nao and human	0.77-1.12
Best communication distance between Nao and human	0.9

Table 1. Performance comparison.

The use of different image processing techniques to recognize the number of fingers was significantly affected by the position of the face, the background, and the luminosity of the area. Therefore, a handheld device-based interaction system was provided for the child to interact with the robot. The device was connected via WiFi to the robot to reduce the restriction of a child's mobility. The projector in this mixed reality system also improved the child's learning.

Conclusion

In this study, the author proposed and tested a mixed reality-based learning toolkit that enhances learning by children and increased the level of intelligence for the robot Nao among 3-7 years old children. The teacher (in this case) and the robot Nao interacted with the child through various means and environments. This child's learning is improved using different alternatives for each process involved. Considering the growing use of handheld devices by children, the designed system architecture of the robot was applied to make learning fun for the child. Therefore, such common problems as the impact of the environment on image processing were avoided. A comparison was also made between the responsiveness of the recognition algorithms in Nao and the control server in this system. The author proposes that the Nao level of intelligence could be increased to a higher level in the future for teaching different subjects to the children, apart from mathematics.

Acknowledgment

The author thanks the Science and Technology Unit at King Abdulaziz University. This project was funded by the National Plan for Science, Technology, and Innovation (MAARIFAH), King Abdulaziz City for Science and Technology, the Kingdom of Saudi Arabia (award number (03-INF188-08)).

References

1. Poole, David. "The Independent Choice Logic for Modelling Multiple Agents under Uncertainty." Arti Intell 94 (1997): 7-56.
2. Solomonoff, Ray J. "The Time Scale of Artificial Intelligence: Reflections on Social Effects." *Human Syst Manag* 5 (1985): 149-153.
3. Pfeifer, Rolf, Iida Fumiya, and Bongard Josh. "New Robotics: Design Principles for Intelligent Systems." Arti Life 11 (2005): 99-120.
4. Chassignol, Maud, Khoroshavin Aleksandr, Klimova Alexandra, and Bilyatdinova Anna. "Artificial Intelligence Trends in Education: A Narrative Overview." Proc Compu Sci 136 (2018): 16-24.
5. Sarama, Julie, and Clements Douglas H. "Design of Microworlds in Mathematics and Science Education." J Edu Compu Res 27 (2002): 1-5.

How to cite this article: Baothman, Fatmah Abdulrahman. "An Artificial Intelligence Design Principle Toolkit for Teaching Basic Mathematics to Children." *Adv Robot Autom*10 (2021) : 199